CONTEMPORARY EVOLUTION IN THREESPINE STICKLEBACK

FROM UPLIFTED ISLANDS IN ALASKA

By

Emily Lescak

RECOMMENDED:

Dr. William Cresko

Dr. Benjamin Harrison

Dr. Megan McPhee

Dr. J. Andres Lopez
Advisory Committee Co-Chair

Dr. Frank von Hippel
Advisory Committee Co-Chair

Dr. Shannon Atkinson
Chair, Graduate Program in Fisheries

APPROVED:

Dr. Joan Braddock
Interim Dean, School of Fisheries and Ocean Sciences

Dr. John Eichelberger
Dean of the Graduate School

27 July 2015
Date

CONTEMPORARY EVOLUTION IN THREESPINE STICKLEBACK FROM UPLIFTED ISLANDS IN ALASKA

A

DISSERTATION

Presented to the Faculty

of the University of Alaska Fairbanks

in Partial Fulfillment of the Requirements

for the Degree of

DOCTOR OF PHILOSOPHY

by

Emily Lescak, M.S.

Fairbanks, AK

August 2015

Abstract

How rapidly can evolution occur in the wild? Threespine stickleback *(Gasterosteus aculeatus*) fish are a prime model organism to address this question because we can study post-glacial populations that are only about 13,000 years old. During this relatively short period of time, oceanic stickleback have repeatedly and independently colonized newly available freshwater habitat, giving rise to resident freshwater populations that exhibit repeated patterns of phenotypic and genetic divergence. However, it is currently unknown whether it actually takes thousands of years for resident freshwater populations to diverge from the ancestral form. To address this question, we have identified phenotypically variable stickleback populations in freshwater sites formed on islands in Prince William Sound and the Gulf of Alaska as a result of uplift from the 1964 Alaska Earthquake. Population genomics analyses support the hypothesis of ongoing independent colonization of freshwater habitats by oceanic ancestors. Despite recurrent gene flow between oceanic and freshwater stickleback, we find that the magnitude of phenotypic and genetic divergence between the ancestral and derived populations is comparable to what has been observed in populations that were founded thousands of years ago. Our data implicate natural selection as the major driver of phenotypic diversification and support the hypothesis that the metapopulation organization of this species helps maintain a large pool of standing genetic variation available for selection when oceanic stickleback colonize fresh water. We propose that the greatest amount of phenotypic evolution occurs within the first few decades after stickleback colonize novel freshwater environments.

Table of Contents

Page

Page

Page

Page

List of Figures

Page

Page

Page

List of Tables

Page

Page

Dedication

For Mia Rowan.

Acknowledgements

I would like to thank my advisors, Andres Lopez and Frank von Hippel, as well as my committee members, Bill Cresko, Ben Harrison, and Megan McPhee, for productive discussions, training opportunities, and feedback. Funding for this research came largely from a National Science Foundation grant awarded to Frank von Hippel and Bill Cresko. Supplemental funding was provided by LGL Limited, the American Fisheries Society, the University of Alaska Center for Global Change and Arctic Systems Research, the University of Alaska Anchorage, Alaska INBRE, and the Rasmuson Fisheries Board.

This work would not have been possible without my very talented co-authors. Susie Bassham made sequencing libraries used in Chapters 2 and 3 and provided valuable edits to Chapter 2. Julian Catchen created the *Stacks* bioinformatics software we used for analysis of genomic data in Chapters 2 and 3. Ofer Gelmond made the initial field collections on Middleton and Montague Islands and collected the morphological measurements used in Chapter 1. University of Alaska Anchorage undergraduates Mary Sherbick, Ryan Lucas, Dan Prince, and Jeff Colgren assisted with phenotypic and genetic data collection in Chapters 2 and 3. Matt Wund also provided phenotype data and analytical support for Chapter 3. Frank von Hippel and Bill Cresko guided the study designs and analytical approaches for all three chapters.

I would also like to thank the members of the Cresko, von Hippel, and Lopez labs for feedback, training, encouragement, and support. Special thanks go to Ann Petersen, Thom Nelson, Kristin Alligood, Conor O'Brien, Mark Currey, Kate Ituarte, Taylor Wilson, Mac Campbell, Thaddaeus Buser, James Willacker, Buck Furin, Dona Eidam, Heidi Weigner, and Kyle Shedd. I would also like to thank Scott Christy for his assistance in selection of field

sites as well as my field assistants, Lauren Smayda, Veronica Padula, Brian Lohman, and Kelly Walton. Scott Hatch also played a vital role in logistics for traveling to and sampling on Middleton Island.

Much gratitude goes to my husband, Peter Bellino, for enduring my intense travel schedule and prolonged absences and continuously providing support and encouragement. Thanks also to my parents for supporting my desires to learn and explore from a young age.

All research was approved by the University of Alaska Anchorage Institutional Animal Care and Use Committee and all field collections were compliant with Alaska Department of Fish and Game permits.

Introduction

Wild populations are capable of evolutionary rates that are orders of magnitude faster than those estimated based on the fossil record (Gingerich, 1983; Reznick *et al.*, 1997; Thompson, 1998; Hendry & Kinnison, 1999). Many examples of contemporary evolution - or changes observable within less than one hundred years (Hendry & Kinnison, 1999; Kinnison & Hendry, 2001) - result from the establishment of new populations (i.e. colonization) or environmental alterations (Reznick & Ghalambor, 2001). Exposure to novel resources, new environments, new predation regimes, and novel competitors are all factors that have been found to induce contemporary change in wild populations (Reznick & Ghalambor, 2001). Accelerated evolutionary rates are often attributed to high selection coefficients of phenotypes that are important in adaptation (Endler, 1986; Reznick & Ghalambor, 2001). When abrupt environmental change elicits strong directional selection on a population, it is predicted that the rate of evolution will initially be high, then decline as the population approaches fitness optima (Orr, 1998).

A challenge in studying contemporary evolution is that it is nearly impossible to predict natural disturbances that might elicit evolutionary change, making identification of study populations and the ancestral state difficult (Hairston & Walton, 1986). Timing, determination of traits undergoing rapid change, and knowledge of the ancestral state are all vital for documenting contemporary evolution, but can easily be missed. Despite these challenges, rapid change has been documented across a wide range of taxa in traits associated with morphology, physiology, life history, phenology, and behavior. The majority of examples of contemporary evolution involve complex traits, which are

influenced by the environment and interactions among multiple genes (Reznick & Ghalambor, 2001).

To demonstrate that changes are due to evolution rather than plasticity, evidence is needed to show that genetic changes have taken place (Merilä & Hendry, 2014). Since little is known about the genetic makeup of most wild populations (Ellegren & Sheldon, 2008; Mackay *et al.*, 2009), it is challenging to conclusively demonstrate that evolution has occurred (Merilä & Hendry, 2014). Few studies have convincingly shown that phenotypic changes on contemporary timescales have an underlying genetic basis (Gienapp *et al.*, 2008; Merilä, 2012). However, advances in sequencing technology are now allowing for detailed investigations of the genetic basis for contemporary change in wild populations. For example, a transcriptomics scan of wild yellow perch (*Perca flavescens*) originating from clean and contaminated lakes revealed a genetic basis for accelerated life-cycle completion in individuals occupying contaminated lakes over a period of about 85 years (Bélanger-Deschênes *et al.*, 2013).

The threespine stickleback (*Gasterosteus aculeatus*) is an ideal candidate model for studying contemporary evolution in the wild. Throughout much of the last century, the stickleback's ecology, behavior, physiology, morphology, and development have been carefully documented throughout its holarctic distribution (see Bell & Foster, 1994 and von Hippel, 2010 for reviews). Oceanic stickleback are the ancestral form and occur as either marine (strictly oceanic) or anadromous (entering freshwater to breed) life history types. Throughout their distribution, they colonize newly available freshwater habitat, where they give rise to resident freshwater populations that exhibit stereotypical phenotypic

divergence from the ancestral state, which includes reductions in body size and external bony armor and changes in body shape (reviewed by Bell & Foster, 1994).

During the relatively short evolutionary time of about 13,000 years (roughly 6,500 generations; Baker *et al.*, 1998), oceanic stickleback colonized newly available freshwater habitat in Alaska's Cook Inlet Basin, giving rise to resident freshwater populations that are both phenotypically and genetically divergent (e.g., Baker *et al.*, 1998; Cresko *et al.*, 2004; Karve *et al.*, 2008; Aguirre, 2009; Chan *et al.*, 2010; Hohenlohe *et al.*, 2010). An open question is whether it actually takes thousands of years for freshwater phenotypes to evolve, or if these changes can occur in much shorter timescales. Studies of lateral plate loss over contemporary time periods in Norway, Iceland, and Alaska following colonization of freshwater habitat by oceanic stickleback (Klepaker, 1993; Kristjansson *et al.*, 2002; Bell *et al.*, 2004; Kristjansson, 2005) suggest that evolution can occur in this species in a matter of decades. However, to date, underlying genomic changes in wild populations over contemporary time scales have not been investigated.

To address questions regarding the tempo of evolution in the wild, we sought a study area with a known approximate age in which oceanic stickleback were the most likely colonizers. We found such a study area in Prince William Sound and the Gulf of Alaska. In 1964, uplift from the largest earthquake ever recorded in North America resulted in the exposure of previously submarine terraces and isolation of new freshwater habitat on multiple islands. Over a period of three field seasons (2005, 2010, and 2011), we collected resident freshwater and oceanic threespine stickleback from three uplifted islands: Montague and Danger Islands, located in Prince William Sound, and Middleton Island, in the Gulf of Alaska. Our first goal was to characterize patterns of phenotypic

variation in resident freshwater and oceanic stickleback. Using reduced-representation sequencing, we generated an extensive dataset of genotypes that we used to test the hypothesis that these resident freshwater populations were derived from oceanic ancestors less than fifty years ago and to quantify the extent of genetic divergence between ancestral oceanic and derived freshwater individuals. Lastly, we addressed the question of whether ancient divergence influences contemporary patterns of phenotypic and genetic variation by testing for relationships between ancestral mitochondrial lineage and both a suite of phenotypes as well as nuclear genetic variation.

References

Aguirre, W. E. 2009. Microgeographical diversification of threespine stickleback: body shape and habitat correlations in a small, ecologically diverse Alaskan drainage. *Biol J Linn Soc* **98**: 139-151.

Baker, J. A., Foster, S. A., Heins, D. C., Bell, M. A. & King, R. W. 1998. Variation in female life-history traits among Alaskan populations of the threespine stickleback, Gasterosteus aculeatus L. (Pisces: Gasterosteidae). *Biol J Linn Soc* **63**: 141-159.

Bélanger-Deschênes, S., Couture, P., Campbell, P. G. & Bernatchez, L. 2013. Evolutionary change driven by metal exposure as revealed by coding SNP genome scan in wild yellow perch (*Perca flavescens*). *Ecotoxicology* **22**: 938-957.

Bell, M. A. & Foster, S. A. 1994. *The evolutionary biology of the threespine stickleback*. Oxford University Press, USA.

Bell, M. A., Aguirre, W. E. & Buck, N. J. 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58**: 814-824.

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Shapiro, M. D., Brady, S. D. et al. 2010. Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. *Science* **327**: 302-305.

Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P. et al. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc Natl Acad Sci USA* **101**: 6050-6055.

Ellegren, H. & Sheldon, B. C. 2008. Genetic basis of fitness differences in natural populations. *Nature* **452**: 169-175.

Endler, J. A. 1986. *Natural selection in the wild*. Princeton University Press, USA.

Gienapp, P., Teplitsky, C., Alho, J. S., Mills, J. A. & Merilä, J. 2008. Climate change and evolution: disentangling environmental and genetic responses. *Mol Ecol* **17**: 167-178.

Gingerich, P. D. 1983. Rates of evolution: effects of time and temporal scaling. *Science* **222**: 159-161.

Hairston, N. G. & Walton, W. E. 1986. Rapid evolution of a life history trait. *Proc Natl Acad Sci U S A* **83**: 4831-4833.

Hendry, A. P. & Kinnison, M. T. 1999. The Pace of Modern Life: Measuring Rates of Contemporary Microevolution. **53**: 1637-1653.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. & Cresko, W. A. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**: e1000862-e1000862.

Karve, A. D., von Hippel, F. A. & Bell, M. A. 2008. Isolation between sympatric anadromous and resident threespine stickleback species in Mud Lake, Alaska. *Environ Biol Fish* **81**: 287-296.

Kinnison, M. T. & Hendry, A. P. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* **112-113**: 145-164.

Klepaker, T. 1993. Morphological changes in a marine population of threespined stickleback, *Gasterosteus aculeatus*, recently isolated in fresh water. *Can J Zool* **71**: 1251-1258.

Kristjansson, B. K. 2005. Rapid morphological changes in threespine stickleback, *Gasterosteus aculeatus*, in freshwater. *Environ Biol Fish* **74**: 357-363.

Kristjansson, B. K., Skúlason, S. & Noakes, D. L. G. 2002. Morphological segregation of Icelandic threespine stickleback (*Gasterosteus aculeatus* L). *Biol J Linn Soc* **76**: 247-257.

Mackay, T. F. C., Stone, E. A. & Ayroles, J. F. 2009. The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* **10**: 565-577.

Merilä, J. 2012. Evolution in response to climate change: in pursuit of the missing evidence. *BioEssays* **34**: 811-818.

Merilä, J. & Hendry, A. P. 2014. Climate change, adaptation, and phenotypic plasticity: the problem and the evidence. *Evol Appl* **7**: 1-14.

Orr, H. A. 1998. The population genetics of adaptation: the distribution of factors fixed during adaptive evolution. *Evolution* **52**: 935-949.

Reznick, D. N. & Ghalambor, C. K. 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* **112-113**: 183-198.

Reznick, D. N., Shaw, F. H., Rodd, F. H. & Shaw, R. G. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* **275**: 1934-1937.

Thompson, J. N. 1998. Rapid evolution as an ecological process. *Trends Ecol Evol* **13**: 329-332.

von Hippel, F. 2010. *Tinbergen's legacy in behaviour: sixty years of landmark stickleback papers*. Brill, Netherlands.

Chapter 1 Contemporary Phenotypic Evolution in Threespine Stickleback from Earthquake-Uplifted Ponds on Middleton Island, Alaska[1]

Abstract

Phenotypic changes can occur rapidly when organisms experience fundamental environmental shifts. Despite recent examples of contemporary evolution - genetic changes on the order of decades or even years – it is still unclear how common this phenomenon occurs in natural populations. We quantified phenotypic divergence of resident freshwater threespine stickleback (*Gasterosteus aculeatus*) from oceanic ancestors over the course of a few decades on Middleton Island, Alaska. Seismic uplift from the 1964 Great Alaska Earthquake caused the formation of a new terrace around the island on which freshwater ponds formed. We studied resident freshwater and sympatric oceanic stickleback from these ponds to determine the extent to which freshwater phenotypes have evolved and the degree to which ecotypic divergence is maintained despite potential for introgressive hybridization. We found that these young resident freshwater populations have undergone coordinated phenotypic divergence from oceanic ancestors in traits associated with feeding, bony armor, swimming, and body size - all heritable traits. Furthermore, we found that while this multivariate ecotypic divergence was replicated across populations in different watersheds, significant morphological differences were also found among some freshwater populations. These data support a hypothesis of rapid and parallel phenotypic divergence after multiple colonization events by oceanic ancestors, and emphasize the important roles of strong natural selection and reproductive isolation in shaping complex

[1] Lescak, E.A., Gelmond, O., von Hippel, F.A., and Cresko, W.A. Prepared for submission in Journal of Evolutionary Biology.

phenotypic diversity. When challenged by a new environment, oceanic stickleback appear to have the ability to evolve on a time scale of decades.

Introduction

Contemporary evolution occurring over a span of decades historically has been considered rare in the wild, with the industrial melanism of *Biston betularia* (Kettlewell, 1958) long serving as a primary and apparently unusual example. While theory has predicted that selection is often a weak and diffuse evolutionary process in natural populations (Lande, 1979; Lande & Arnold, 1983), artificial selection can clearly produce rapid evolutionary responses. However, artificial selection usually targets a small number of traits, whereas natural selection acts on the myriad complex phenotypes that contribute to fitness in nature. Empirical studies in the wild have demonstrated that strong selection and contemporary evolution may be relatively common (Endler, 1986; Schluter, 2000a; Schluter, 2000b; Grant & Grant, 2002). In addition to being of fundamental importance for understanding how diversification occurs in the wild, better documentation of this phenomemon is of practical importance because it may play a significant role in species survival in the face of global change (Endler, 1986; Stockwell & Weeks, 1999; Reznick & Ghalambor, 2001; Gienapp *et al.*, 2008).

The ability of a species to evolve is influenced by a host of factors including population size, genetic diversity and the strengths of selection and gene flow (Wright, 1931; Kimura, 1983). Both empirical data and theory predict a high likelihood of contemporary evolution in scenarios in which trait heritability is high and the intensity of directional selection is strong (Stockwell *et al.*, 2003). Intense directional selection is predicted to be strongest, and therefore evolution most rapid, immediately after an

environmental disturbance, with the pace predicted to gradually decrease as the population approaches a fitness optimum (Endler, 1986; Hendry & Kinnison, 1999; Kinnison & Hendry, 2001; Orr, 2005).

Examples of contemporary evolution have been cited in the literature across a small but growing range of taxa (e.g., Harvey, 1990; Losos *et al.*, 1997; Hendry & Kinnison, 1999; Lee, 1999; Huey *et al.*, 2000; Bone & Farres, 2002; Reznick & Ghalambor, 2001; Grant & Grant, 2002; Stockwell *et al.*, 2003; Svensson & Gosden, 2007), with many of the best examples coming from wild fish populations (e.g., Reznick *et al.*, 1997; Stockwell & Weeks, 1999; Hendry *et al.*, 2000; Kinnison *et al.*, 2001; Quinn *et al.*, 2001; Koskinen *et al.*, 2002; O'Steen *et al.*, 2002). In most of these systems, traits have a known genetic basis and in many cases, divergence occurred after colonization of new habitat or abrupt changes to the environment (Losos *et al.*, 1997; Thompson, 1998; Hendry & Kinnison, 1999; Lee, 1999; Huey *et al.*, 2000; Stockwell *et al.*, 2003). However, examples of replicated patterns of divergence are few and the genetic and genomic basis of contemporary evolution remains unclear. Knowing the genetic basis of divergence will facilitate a better understanding of the mechanisms governing rapid evolutionary responses. Therefore, documenting contemporary phenotypic evolution in an organism with significant genomic resources (e.g. annotated reference genome, developed and refined genetic tools) is a key first step in addressing this fundamental evolutionary problem.

A promising system for testing hypotheses regarding contemporary evolution is the threespine stickleback fish (*Gasterosteus aculeatus*), which shows marked evolutionary divergence between ancestral oceanic and derived freshwater forms (e.g., Klepaker, 1993; Bell, 2001; Kristjansson *et al.*, 2002a; Kristjansson *et al.* 2002b; Bell *et al.*, 2004;

Kristjansson, 2005; Kitano *et al.*, 2008; Aguirre & Bell, 2012). The high heritability of adaptive phenotypes and strong selection after colonization of freshwater habitats allow for rapid phenotypic divergence between oceanic and freshwater ecotypes (Klepaker, 1993; Bell, 2001; Kristjansson *et al.*, 2002a; Kristjansson et al. 2002b; Bell *et al.*, 2004; Kristjansson, 2005; Kitano *et al.*, 2008; Aguirre & Bell, 2012).

Over the last 13,000-20,000 years, stickleback populations from northern latitudes have undergone a postglacial adaptive radiation, during which oceanic stickleback have repeatedly and independently colonized newly available freshwater habitats, giving rise to phenotypically divergent populations that demonstrate parallel responses to natural selection (Hagen, 1967; McPhail, 1969; Schluter & McPhail, 1992; Bell & Foster, 1994; Nagel & Schluter, 1998; Schluter, 2000a; Schluter, 2000b; Rundle *et al.*, 2000; McKinnon & Rundle, 2002; Kingsley *et al.*, 2004; Boughman *et al.*, 2005; Boughman, 2007; Cresko *et al.*, 2004, 2007; Hohenlohe *et al.*, 2010; Jones *et al.*, 2012; Kaeuffer *et al.*, 2012). Despite this divergence, gene flow from freshwater into oceanic populations seems to be prevalent (Hohenlohe *et al.*, 2010; Jones *et al.*, 2012).

In addition to the many studies of stickleback divergence documenting evolution at the scale of tens of thousands of years, a small number of studies have revealed patterns of contemporary evolution in traits such as body size and shape, lateral plate count, and gill raker number (Klepaker, 1993; Bell, 2001; Kristjansson *et al.*, 2002; Bell *et al.*, 2004; Kristjansson, 2005; Kitano *et al.*, 2008; Gelmond et al., 2009; Aguirre & Bell, 2012). These studies raise the possibility that most stickleback phenotypic evolution occurs in the first few decades after colonization. Research on additional stickleback populations undergoing contemporary evolution, especially those doing so in parallel on a local scale, will allow

determination of the timing and repeatability of phenotypic change, and ultimately, its genetic basis.

Here we describe five populations of resident freshwater threespine stickleback in the very early stages of divergence from oceanic ancestors on Middleton Island, located in the Gulf of Alaska. These populations are found in coastal seismically uplifted ponds that were formed as a result of the 1964 Great Alaska Earthquake. We previously reported morphological divergence between freshwater and oceanic stickleback on Middleton Island on a trait-by-trait basis (Gelmond *et al.*, 2009). In this paper, we document the pattern of multivariate phenotypic divergence in numerous traits indicative of adaptation to freshwater habitats and test whether this divergence is maintained despite potential for hybridization in ponds containing sympatric resident freshwater and oceanic ecotypes.

We find that in less than fifty years, freshwater stickleback have diverged from oceanic ancestors across a suite of trophic, bony armor, swimming, and body size traits in magnitude similar to what has been documented for stickleback populations formed post-glacially (~13,000 ybp; Bell & Foster, 1994). Although divergence was replicated across populations in different watersheds, suggesting a parallel nature to this rapid adaptation, significant morphological differences were also found among some freshwater populations. Our results support a hypothesis of multiple colonization events by oceanic ancestors and potentially important roles for geographic isolation, strong natural selection, genetic drift, and reproductive isolation in shaping complex phenotypic diversity.

Methods

Study site

Middleton Island (59.435967°N, 146.323875°W) is located at the edge of the continental shelf in the Gulf of Alaska (Fig. 1.1). The island is 8 km long, 1.6 km wide, and its maximum elevation is 47.6 m. One obvious feature is its step-like character, derived from a well-preserved sequence of tectonically uplifted marine terraces (Plafker & Rubin, 1978). On March 27, 1964, the second most powerful earthquake ever recorded in the world, and the largest ever in North America, struck south-central Alaska. Registering 9.2 on the Richter scale, this catastrophic event caused Middleton Island to experience 3.4 m of immediate vertical uplift, which created a new terrace around the margins (Plafker, 1967, 1969; Plafker & Rubin, 1978). Other than two sites on pre-1964 terraces, ponds found on the post-1964 terrace are the only year-round water bodies on the island (Fig. 1.1).

Stickleback collection and preservation

Stickleback were collected from six sites on the post-1964 terrace (Fig. 1.1) during May 25 – 31 and July 2 – 5, 2005 using 0.32 and 0.64 cm mesh unbaited minnow traps left near shore overnight. Salinity was measured using a YSI-30 meter. Five sites were classified as freshwater habitat (salinity ranges from 0.17-0.72 ppt), while one was classified as oceanic habitat (salinity = 21.36 ppt). Stickleback were euthanized with an overdose of neutral pH MS-222 anesthetic, fixed in 10% neutral buffered formalin, bleached with hydrogen peroxide, stained with Alizarin red S, and preserved in 70% ethanol.

Morphological analyses

Stained fish were photographed using a tripod-mounted Canon digital SLR camera and Canon EOS ViewerUtility software. Only fish longer than 32 mm standard length (SL; anterior tip of upper jaw to posterior end of hyperal plate) were analyzed such that only adults were included (Bell, 2001). The lengths of the left pelvic spine and second dorsal spine were measured using digital calipers. The number of lateral plates on the left side and number of gill rakers on the first right arch were counted under a dissecting microscope. Gill raker number was recorded from ≤ 30 fish per site.

The following traits were measured digitally to the nearest 0.01 mm from photographs using tpsDig2 v.2.04 software: SL, snout length (anterior tip of upper jaw to closest margin of orbit), vertical eye diameter, horizontal operculum width, length of pectoral fin insertion, length of the first lateral plate behind the second dorsal spine, distance between the insertion of the second and third dorsal spines, and length of the base of the dorsal fin (Fig. 1.2). The numbers of principal anal, dorsal, and caudal fin rays were counted. Also recorded was the diagonal length of the smooth area anterior to the left pectoral fin, as a proxy of the size of the fin abductor muscle.

Statistical analyses

Freshwater and oceanic stickleback differ markedly in size, defensive armor and swimming traits (Bell & Foster, 1994; von Hippel & Weigner, 2004; Aguirre *et al.*, 2008; Karve *et al.*, 2008; Gelmond *et al.*, 2009). Each individual was visually classified as being either oceanic or freshwater based on these features. To verify the visual classification, K-Means clustering was performed using five traits: SL, number of left lateral plates, and size-adjusted lengths of the left pelvic spine, second dorsal spine, and smooth diagonal surface

anterior to the pectoral fin. Statistical analyses were concordant with prior visual classifications: only 1.2% of the individuals were categorized differently by the cluster analysis than they were by visual classification. For all subsequent statistical tests, the results of the visual classification were used to assign each fish to its category.

Principal Components Analyses (PCA) were used to uncover the major axes of phenotypic variation among 557 individuals from six populations (Table 1.1). Not all traits were reliably visible on all photos and gill rakers were not counted on all individuals; to accommodate these missing morphological data, the ppca function in the pcaMethods package was used to perform probabilistic PCAs that impute missing data (Stacklies et al., 2007). Since morphometric traits exhibit allometric growth patterns, we performed PCAs with size-standardized data. To size-standardize the data, we used a Model I regression to test for a significant relationship between each morphometric trait and SL separately for each population. For those traits with significant regressions, the residuals were calculated and included in PCAs along with unstandardized data for meristic traits and morphometric traits that did not have significant size relationships.

Because principal components (PCs) are by definition orthogonal, we analyzed each separately. Separate ANOVAs were performed for each of the first three PCs, followed by Tukey post-hoc tests, to determine whether populations and ecotypes significantly differed in morphology. For population-level analyses, oceanic individuals from sympatric sites (8, 9, 13, and 19) were pooled due to small sample sizes and compared to each of five freshwater populations as well as oceanic individuals from the marine habitat. All analyses were performed in R (R Core Team, 2013).

Results

Freshwater and oceanic stickleback occured in sympatry. The oceanic site (site 23) contained only individuals with oceanic phenotypes (Table 1.1). Of the five freshwater sites, one contained only individuals with freshwater phenotypes (site 6) and four contained both freshwater and oceanic individuals (sites 8, 9, 13, 19). In these sympatric sites, oceanic individuals comprised between 2% and 62% of the fish (Table 1.1).

Freshwater and oceanic stickleback were phenotypically divergent. The two ecotypes differed significantly in standard length ($F_{1,569}$ = 943.80; P < 0.001; Fig. 1.3, Table 1.1) and lateral plate number ($F_{1,569}$ = 1,734.00; P < 0.001; Fig. 1.3; Table 1.1). When using a multivariate analysis of all data, the first principal component (PC 1) explained 93% of the total variation and clearly separated freshwater and oceanic individuals (Table 1.2, Fig. 1.4). Ecotypes differed significantly in scores from PC 1 ($F_{1,553}$ = 786.52, P < 0.001), PC 2 ($F_{1,553}$ = 16.11, P < 0.001), and PC 3 ($F_{1,553}$ = 14.33, P < 0.001; Fig. 1.5A).

Because lateral plates are one of the most easily identifiable differences between oceanic and freshwater stickleback, and they had the largest loadings on PC 1 (Table 1.1), we removed the lateral plate variable from the dataset to determine the extent to which other traits contributed to divergence between ecotypes. The first PC from this new analysis without lateral plates accounted for 28% of the variation, while PC 2 accounted for 16% (Table 1.2; Fig. 1.4). Even without the inclusion of lateral plates, the two ecotypes differed significantly in scores from PC 1 ($F_{1,546}$ = 95.30, P < 0.001), PC 2 ($F_{1,546}$ = 12.09, P < 0.001), and PC 3 ($F_{1,546}$ = 4.23, P = 0.040; Fig. 1.5B), indicating that traits other than lateral plates contribute to phenotypic divergence between oceanic and freshwater ecotypes. Lengths of the left pelvic spine, second dorsal spine, proxy for the pectoral muscle, and

pectoral fin insertion had the largest loadings on PC 1, while snout, eye, and operculum lengths as well as the distance between dorsal spines had the largest loadings on PC 2 (Table 1.1).

Freshwater populations were morphologically variable. In addition to differences between oceanic and freshwater ecotypes, we also tested for differences among populations, including all freshwater populations, all oceanic fish from freshwater sites pooled into one population, and the single oceanic population in the oceanic site. In the analysis including lateral plates, populations significantly differed in scores for PC 1 ($F_{6,550}$ = 275.00, $P < 0.001$), PC 2 ($F_{6,550}$ = 4.02, $P < 0.001$), and PC 3 ($F_{6,550}$ = 8.68, $P \leq 0.001$). PC 1 grouped the freshwater populations separately from the oceanic populations and also separated freshwater site 9 from sites 8 and 19 (Fig. 1.5C; Table 1.2). Significant differences occurred among some freshwater populations in scores from PC 2 and PC 3 and PC 3 also separated the two oceanic groups.

When lateral plates were excluded, populations still significantly differed in scores from PC 1 ($F_{6,550}$ = 27.88, $P < 0.001$), PC 2 ($F_{6,550}$ = 9.93, $P \leq 0.001$) and PC 3 ($F_{6,550}$ = 7.38, $P < 0.001$). PC 1 again separated the freshwater and oceanic sites, while also separating sites 9 and 13 (Fig. 1.5D; Table 1.2). Significant differences also occurred among some freshwater populations in scores from PC 2 and PC3 and PC 2 separated the two oceanic groups.

Discussion

Freshwater and oceanic stickleback were differentiated across multiple morphological features. As expected, differences in SL and lateral plate count (Fig. 1.3; Table 1.1) were the dominant features that differ. However, even when these factors were removed, traits associated with foraging (e.g., gill rakers), defense (e.g., spines) and

swimming (e.g., pectoral muscle and fin) also distinguish the two ecotypes, as has been found in previous studies (Table 1.1; Bell & Foster, 1994; von Hippel & Weigner, 2004; Aguirre *et al.*, 2008; Karve *et al.*, 2008; Gelmond *et al.*, 2009). Furthermore, many of these traits have a known genetic basis (Colosimo *et al.*, 2004, 2005; Cresko *et al.*, 2004). The evolution of multiple phenotypes that have been shown to be genetically determined and ecologically important in other stickleback populations, including those in Alaska, strongly suggests that the phenotypic divergence we observe on Middleton Island has a large heritable component.

Even in freshwater habitats where oceanic and freshwater ecotypes were sympatric, they were phenotypically divergent despite potential for hybridization. Oceanic fish from these sites demonstrate significant differences in phenotype when compared to oceanic fish found in marine habitat (Fig. 1.5, Table 1.2), suggesting that at least some of the oceanic fish in freshwater habitat are either freshwater-oceanic hybrids or in the early stages of divergence from recolonized marine ancestors. The newly formed freshwater habitats may not be completely isolated from the ocean, suggesting that secondary contact between oceanic and freshwater fish has likely been repeated over many years. This scenario creates potential for recolonization of freshwater sites by marine fish. Despite the possibility for introgressive hybridization, oceanic and freshwater ecotypes in these freshwater habitats were strongly differentiated from one another (Figs. 1.3-1.5; Table 1.2).

Discontinuous trait distributions between ecotypes are likely due to strong divergent selection to the alternative oceanic and freshwater habitats, which may be reinforced by reproductive isolating mechanisms (McPhail, 1994). The strength of selection

on the *Eda* locus in oceanic stickleback transplanted to freshwater ponds is high (0.52; Barrett *et al.*, 2009), which would explain the rapid loss of lateral plates after colonization of fresh water, and assortative mating can evolve in freshwater stickleback populations within a couple of decades of divergence from oceanic ancestors (Furin *et al.*, 2012).

While differentiation over just a few decades is a rate that is seldom observed in nature (reviewed by Hendry *et al.*, 2007), it has previously been documented in threespine stickleback (Bell, 2001; Kristjannson *et al.*, 2002a; Kristjannson *et al.* 2002b; Bell *et al.*, 2004; Kristjannson, 2005; Aguirre *et al.*, 2008) as well as other fish species. For example, ancestral and derived populations of mosquitofish (*Gambusia affinis*) diverged in several life history traits in less than 60 years (Stockwell & Weeks, 1999). Similarly, Pacific salmon (*Oncorhynchus* spp.) adapted to different breeding environments and developed partial reproductive isolation over a period of 13-26 generations (Hendry *et al.*, 2000; Kinnison *et al.*, 2001; Quinn *et al.*, 2001). Parallel life history evolution has also been observed in the guppy (*Poecilia reticulata*) in only 4-11 years as a result of changes in predation pressure (Reznick *et al.*, 1997).

Rapid evolution of Middleton Island stickleback is especially remarkable because it appears to have happened in different watersheds on the island. Patterns of phenotypic variation support the hypothesis of multiple, independent colonization events by oceanic ancestors. Four sites from different watersheds on the 1964 terrace contain individuals with oceanic and freshwater phenotypes in sympatry (Fig. 1.1; Table 1.1), consistent with the well-described ability of oceanic stickleback to colonize freshwater habitats (Bell & Foster, 1994).

How does phenotypic divergence across a suite of traits evolve over a period of decades? Previous studies of stickleback throughout their holarctic distribution suggest that rapid evolution in fresh water is based on standing genetic variation in the oceanic populations, rather than the acquisition and fixation of novel mutations. Freshwater alleles are maintained at low frequency in oceanic populations due to low levels of hybridization between the two ecotypes. Since the alleles for freshwater phenotypes are already present in oceanic colonizers, they are able to quickly rise in frequency upon invasion of fresh water (Schluter & Conte, 2009). For example, the same alleles of *Eda* and *Pitx1* underlie lateral plate and pelvic girdle variation, respectively, in geographically disparate populations (Colosimo *et al.*, 2004, 2005; Cresko *et al.*, 2004; Chan *et al.*, 2010; Jones *et al.*, 2012). Similarly, independently colonized freshwater populations in Cook Inlet, Alaska show remarkably similar signatures of selection when they are compared to ancestral oceanic populations, suggesting that the same genomic locations are repeatedly differentiated in freshwater populations (Hohenlohe *et al.*, 2010).

Phenotypic divergence may also be accelerated due to the genomic architecture of stickleback. By reducing recombination, chromosomal inversions may allow for rapid fixation of adaptive phenotypes relating to osmoregulation, bony armor, and body shape (Jones *et al.*, 2012; Feulner *et al.*, 2013) and long- and short-range linkage disequilibrium may facilitate rapid evolution of armor loss in freshwater populations by keeping alleles for freshwater phenotypes in association (Hohenlohe *et al.*, 2012). A continued area of active research will focus on the genetic and genomic architecture underlying the rapid evolution of stickleback such as that seen in Middleton Island populations.

We have documented that within less than a five-decade (approximate 25 generation; Baker *et al*., 1998) period, resident freshwater stickleback on Middleton Island have shown a replicated pattern of phenotypic differentiation from oceanic stickleback. Surprisingly, the extent of phenotypic divergence is of the same magnitude as what has been described in populations that are thousands of years old (e.g., Bell & Foster, 1994; Cresko *et al.*, 2004; Colosimo *et al.*, 2005). Our findings raise the possibility that phenotypic divergence in stickleback occurs on decadal time scales, rather than over thousands of years, before quickly reaching a new equilibrium (von Hippel & Weigner, 2004). This may be the norm rather than the exception in newly formed freshwater stickleback populations. The well-defined, extremely short time frame between the formation of the sites and the emergence of resident freshwater populations on Middleton Island provides an excellent opportunity to study the initial stages of divergence during adaptation to novel environments. To the extent that differentiated phenotypes contribute to pre- and post-zygotic isolation, these populations form an ideal resource for studying the initial stages of speciation as well.

While phenotypic divergence on contemporary timescales has been described in wild populations across a range of taxa, few of these studies include genetic data. The exceptions have been limited to a handful of genetic markers (e.g., Lee, 1999; Koskinen *et al.*, 2002), making it difficult to demonstrate conclusively that contemporary phenotypic change is due to evolution, rather than plasticity, and to test competing hypotheses of colonization history. Genome-wide datasets are necessary to allow precise determination of population parameters as well as identification of genomic regions under selection. Our results are consistent with the hypothesis of multiple, independent colonization events on

Middleton Island, but testing this hypothesis will require a dense genotypic dataset to identify the genomic regions associated with phenotypic divergence (Hohenlohe *et al.*, 2010; Jones *et al.*, 2012).

Acknowledgements

N.A. Bargmann, B.M Guzzetti, and S.A. Hatch assisted with fieldwork logistics and data collection. G. Plafker of the U.S. Geological Survey generously provided records of his work on Middleton Island. The Chugach National Forest, Bureau of Land Management, Federal Aviation Administration, Anchorage Museum of History and Art, and Army Corps of Engineers provided access to their imagery collections. We thank M.A. Bell, R.R. Bernhardt, S.M. Christy, K.J. Crossen, M. Currey, D. Eidam, C.G. Furin, A.M. Gardell, B.R. Harrison, S.B. Haught, P.A. Heiser, J.A. Lopez, M.V. McPhee, J.S. Moore, M. Nishkian, J.H. Postlethwait, K.R. Shedd, L.C. Smayda, and H.L. Weigner for discussions, comments and assistance. Fish were collected under Alaska Department of Fish and Game permit number SF-2005-020, and all research protocols were approved by the University of Alaska Anchorage Institutional Animal Care and Use Committee. This research was supported by a University of Alaska Anchorage Faculty Development Grant and National Science Foundation grant DEB 0919234 to FAvH, National Science Foundation grants DEB 0949053 & IOS 102728 to WAC, and funds to EAL from the University of Alaska Center for Global Change and Arctic Systems Research, the University of Alaska Anchorage, LGL Limited, and the NSF Alaska EPSCoR Landscape Genetics program.

References

Aguirre, W.E. & Bell, M.A. 2012. Twenty years of body shape evolution in a threespine stickleback population adapting to a lake environment. *Biol. J. Linn. Soc.* **105**: 817-831.

Aguirre, W.E., Ellis, K.E., Kusenda, M. & Bell, M.A. 2008. Phenotypic variation and sexual dimorphism in anadromous threespine stickleback: implications for postglacial adaptive radiation. *Biol. J. Linn. Soc.* **95**: 465-478.

Baker, J.A., Foster, S.A., Heins, D.C., Bell, M.A. & King, R.W. 1998. Variation in female life-history traits among Alaskan populations of the threespine stickleback, *Gasterosteus aculeatus* L. (Pisces: Gasterosteidae). *Biol. J. Linn. Soc.* **63**: 141-159.

Barrett, R.D.H., Vines, T.H., Bystriansky, J.S. & Schulte, P.M. 2009. Should I stay or should I go? The *Ectodysplasin* locus is associated with behavioural differences in threespine stickleback. *Biology Letters* **5**: 788-791.

Bell, M.A. 2001. Lateral plate evolution in the threespine stickleback: getting nowhere fast. *Genetics* **112-113**: 445-461.

Bell, M.A., Aguirre, W.E., Buck, N.J., & Wainwright, P. 2004. Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution* **58**: 814-824.

Bell, M.A. & Foster, S.A. 1994. *The evolutionary biology of the threespine stickleback*. Oxford University Press, USA.

Bone, E. & Farres, A. 2001. Trends and rates of microevolution in plants. Trends and rates of microevolution in plants. *Genetica* 112-113:165-182.

Boughman, J.W. 2007. Condition-dependent expression of red colour differs between stickleback species. *J. Evol. Biol.* **20**: 1577-1590.

Boughman, J.W., Rundle, H.D. & Schluter, D. 2005. Parallel evolution of sexual isolation in sticklebacks. *Evolution* **59**: 361-373.

Chan, Y.F., Marks, M.E., Jones, F.C., Villarreal, G., Shapiro, M.D., Brady, S.D. *et al*. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**: 302-305.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, J., Guadalupe, Dickson, M., Grimwood, J. *et al*. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of *Ectodysplasin* alleles. *Science* **307**: 1928-1933.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. *et al*. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology* **2**: e109.

Cresko, W.A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P. *et al*. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *PNAS* **101**: 6050-6055.

Cresko, W.A., McGuigan, K.L., Phillips, P.C. & Postlethwait, J.H. 2007. Studies of threespine stickleback developmental evolution: progress and promise. *Genetica* **129**: 105-126.

Endler, J.A. 1986. *Natural Selection in the Wild*. Princeton University Press, Princeton.

Feulner, P.G.D., Chain, F.J.J., Panchal, M., Eizaguirre, C., Kalbe, N., Lenz, T.L., *et al.* 2013. Genome-wide patterns of standing genetic variation in a marine population of three-spined stickleback. *Mol. Ecol.* **22**: 635-649.

Furin, C.G., von Hippel, F.A. & Bell, M.A. 2012. Partial reproductive isolation of a recently derived resident-freshwater population of threespine stickleback (*Gasterosteus aculeatus*) from its putative anadromous ancestor. *Evolution* **66**: 3277-3286.

Gelmond, O., Von Hippel, F.A. & Christy, M.S. 2009. Rapid ecological speciation in three-spined stickleback *Gasterosteus aculeatus* from Middleton Island, Alaska: the roles of selection and geographic isolation. *J. Fish. Biol.* **75**: 2037-2051.

Gienapp, P., Teplitsky, C., Alho, J.S., Mills, J.A. & Merilä, J. 2008. Climate change and evolution: disentangling environmental and genetic responses. *Mol. Ecol.* **17**: 167-178.

Grant, P.R. & Grant, B.R. 2002. Unpredictable evolution in a 30-year study of Darwin's finches. *Science* **296**: 707-711.

Hagen, D.W. 1967. Isolating mechanisms in threespine sticklebacks (*Gasterosteus*). *J. Fish. Res. Board Can.* **24**: 1637-1692.

Harvey, A.W. 1990. Sexual differences in contemporary selection acting on size in the hermit crab *Clibanarius digueti*. *Am. Nat.* **136**: 292-304.

Hendry, A.P. & Kinnison, M.T. 1999. Perspective: the pace of modern life: measuring rates of contemporary microevolution. *Evolution* **53**: 1637-1653.

Hendry, A.P., Nosil, P. & Rieseberg, L.H. 2007. The speed of ecological speciation. *Funct. Ecol.* **21**: 455-464.

Hendry, A.P., Wenburg, J.K., Bentzen, P., Volk, E.C. & Quinn, T.P. 2000. Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* **290**: 516-518.

Hohenlohe, P.A., Bassham, S., Currey, M. & Cresko, W.A. 2012. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Philos T Roy Soc B: Biological Sciences* **367**: 395-408.

Hohenlohe, P.A., Bassham, S., Etter, P.D., Stiffler, N., Johnson, E.A. & Cresko, W.A. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**: e1000862.

Huey, R.B., Gilchrist, G.W., Carlson, M.L., Berrigan, D. & Serra, L. 2000. Rapid evolution of a geographic cline in size in an introduced fly. *Science* **287**: 308-309.

Jones, F.C., Grabherr, M.G., Chan, Y.F., Russell, P., Mauceli, E., Johnson, J. *et al.* 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**: 55-61.

Kaeuffer, R., Peichel, C.L., Bolnick, D.I. & Hendry, A.P. 2012. Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution* **66**: 402-418.

Karve, A.D., von Hippel, F.A. & Bell, M.A. 2008. Isolation between sympatric anadromous and resident threespine stickleback species in Mud Lake, Alaska. *Environ. Biol. Fish.* **81**: 287-296.

Kettlewell, H.B.D. 1958. A survey of the frequencies of *Biston betularia* (L.) (Lep.) and its melanic forms in Britain. *Heredity* **12**: 51-72.

Kimura, M. 1983. *The neutral theory of molecular evolution.* Cambridge University Press.

Kingsley, D.M., Zhu, B., Osoegawa, K., De, J., Pieter J., Schein, J., Marra, M. *et al.* 2004. New genomic tools for molecular studies of evolutionary change in threespine sticklebacks. *Behaviour* **141**: 1331-1344.

Kinnison, M.T. & Hendry, A.P. 2001. The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* **112-113**: 145-164.

Kinnison, M.T., Unwin, M.J., Hendry, A.P., & Quinn, T.P. 2001. Migratory costs and the evolution of egg size and number in introduced and indigenous salmon populations. *Evolution* **55**: 1656-1667.

Kitano, J., Bolnick, D.I., Beauchamp, D.A., Mazur, M.M., Mori, S., Nakano, T. *et al*. 2008. Reverse evolution of armor plates in the threespine stickleback. *Curr. Biol.* **18**: 769-774.

Klepaker, T. 1993. Morphological changes in a marine population of threespined stickleback, *Gasterosteus aculeatus*, recently isolated in fresh water. *Can. J. Zool.* **71**: 1251-1258.

Koskinen, M.T., Haugen, T.O. & Primmer, C.R. 2002. Contemporary fisherian life-history evolution in small salmonid populations. *Nature* **419**: 826-830.

Kristjansson, B.K. 2005. Rapid morphological changes in threespine stickleback, *Gasterosteus aculeatus*, in freshwater. *Environ. Biol. Fish.* **74**: 357-363.

Kristjansson, B.K., Skulason, S. & Noakes, D.L.G. 2002a. Morphological segregation of Icelandic threespine stickleback (*Gasterosteus aculeatus* L.). *Biol. J. Linn. Soc.* **76**: 247-257.

Kristjansson, B.K., Skulason, S. & Noakes, D.L.G. 2002b. Rapid divergence in a recently isolated population of threespine stickleback (*Gasterosteus aculeatus* L.). *Evol. Ecol. Res.* **4**: 659-672.

Lande, R. 1979. Quantitative genetic analysis of multivariate evolution, applied to brain: body size allometry. *Evolution* **33**: 402-416.

Lande, R. & Arnold, S.J. 1983. The measurement of selection on correlated characters. *Evolution* **37**: 1210-1226.

Lee, C.E. 1999. Rapid and repeated invasions of fresh water by the copepod *Eurytemora affinis*. *Evolution* **53**: 1423-1434.

Losos, J.B., Warheit, K.I. & Schoener, T.W. 1997. Adaptive differentiation following experimental island colonization in Anolis lizards. *Nature* **387**: 70-73.

McKinnon, J.S. & Rundle, H.D. 2002. Speciation in nature: The threespine stickleback model systems. *TREE* **17**: 480-488.

McPhail, J.D. 1969. Predation and the evolution of a stickleback (Gasterosteus). *J. Fish. Board Can.* **26**: 3183-3208.

McPhail, J.D. 1994. Speciation and the evolution of reproductive isolation in the sticklebacks (*Gasterosteus*) of south-western British Columbia. In: *The evolutionary biology of the threespine stickleback* (M.A. Bell & S.A. Foster, eds), pp. 399-437. Oxford University Press, Oxford.

Nagel, L. & Schluter, D. 1998. Body size, natural selection, and speciation in sticklebacks. *Evolution* **52**: 209-218.

O'Steen, S., Cullum, A.J. & Bennett, A.F. 2002. Rapid evolution of escape ability in Trinidadian guppies (*Poecilia reticulata*). *Evolution* **56**: 776-784.

Plafker, G. 1967. The Alaska Earthquake, March 27, 1964: regional effects – surface faults on Montague Island. *US Geol. Surv. Professional Paper* **543-G**:1-42.

Plafker, G. 1969. The Alaska Earthquake, March 27, 1964: regional effects – the tectonics of the March 27, 1964 Alaska Earthquake. *US Geol. Surv. Professional Paper* **543-I**: 1-74.

Plafker, G. & Rubin, M. 1978. Uplift history and earthquake recurrence as deduced from marine terraces on Middleton Island, Alaska. *US Geol. Surv. Open File Rep.* **78**: 687-721.

Quinn, T.P., Hendry, A.P. & Buck, G.B. 2001. Balancing natural and sexual selection in sockeye salmon: interactions between body size, reproductive opportunity and vulnerability to predation by bears. *Evol. Ecol. Res.* **3**: 917-937.

R Core Team. 2013. R: A language and environment for statistical computing. R Foundation for Statistical Computing.

Reznick, D.N. & Ghalambor, C.K. 2001. The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* **112**: 183-198.

Reznick, D.N., Shaw, F.H., Rodd, F.H. & Shaw, R.G. 1997. Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*). *Science* **275**: 1934-1937.

Rundle, H.D., Nagel, L., Boughman, J.W. & Schluter, D. 2000. Natural selection and parallel speciation in sympatric sticklebacks. *Science* **287**: 306-308.

Schluter, D. 2000a. *The Ecology of Adaptive Radiation*. Oxford University Press, Oxford.

Schluter, D. 2000b. Ecological character displacement in adaptive radiation. *Am. Nat.* **156**: S4-S16.

Schluter, D. & Conte, G.L. 2009. Genetics and ecological speciation. *Proc. Natl. Acad. Sci.* **106**: 9955–9962.

Schluter, D. & McPhail, J.D. 1992. Ecological character displacement and speciation in sticklebacks. *Am. Nat.* **140**: 85-108.

Stacklies, W., Redestig, H., Scholz, M., Walther, D. & Selbig, J. 2007. pcaMethods—a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**: 1164-1167.

Stockwell, C.A. & Weeks, S.C. 1999. Translocations and rapid evolutionary responses in recently established populations of western mosquitofish. *Anim. Conserv.* **2**: 103-110.

Stockwell, C.A., Hendry, A.P. & Kinnison, M.T. 2003. Contemporary evolution meets conservation biology. *TREE* **18**: 94-101.

Svensson, E.I. & Gosden, T.P. 2007. Contemporary evolution of secondary sexual traits in the wild. *Funct. Ecol.* **21**: 422-433.

Thompson, J.N. 1998. Rapid evolution as an ecological process. *TREE* **13**: 329-332.

von Hippel, F.A. & Weigner, H. 2004. Sympatric anadromous-resident pairs of threespine stickleback species in young lakes and streams at Bering Glacier, Alaska. *Behaviour* **141**: 1441-1464.

Wright, S. 1931. Evolution in Mendelian populations. *Genetics* **16**: 97-159.

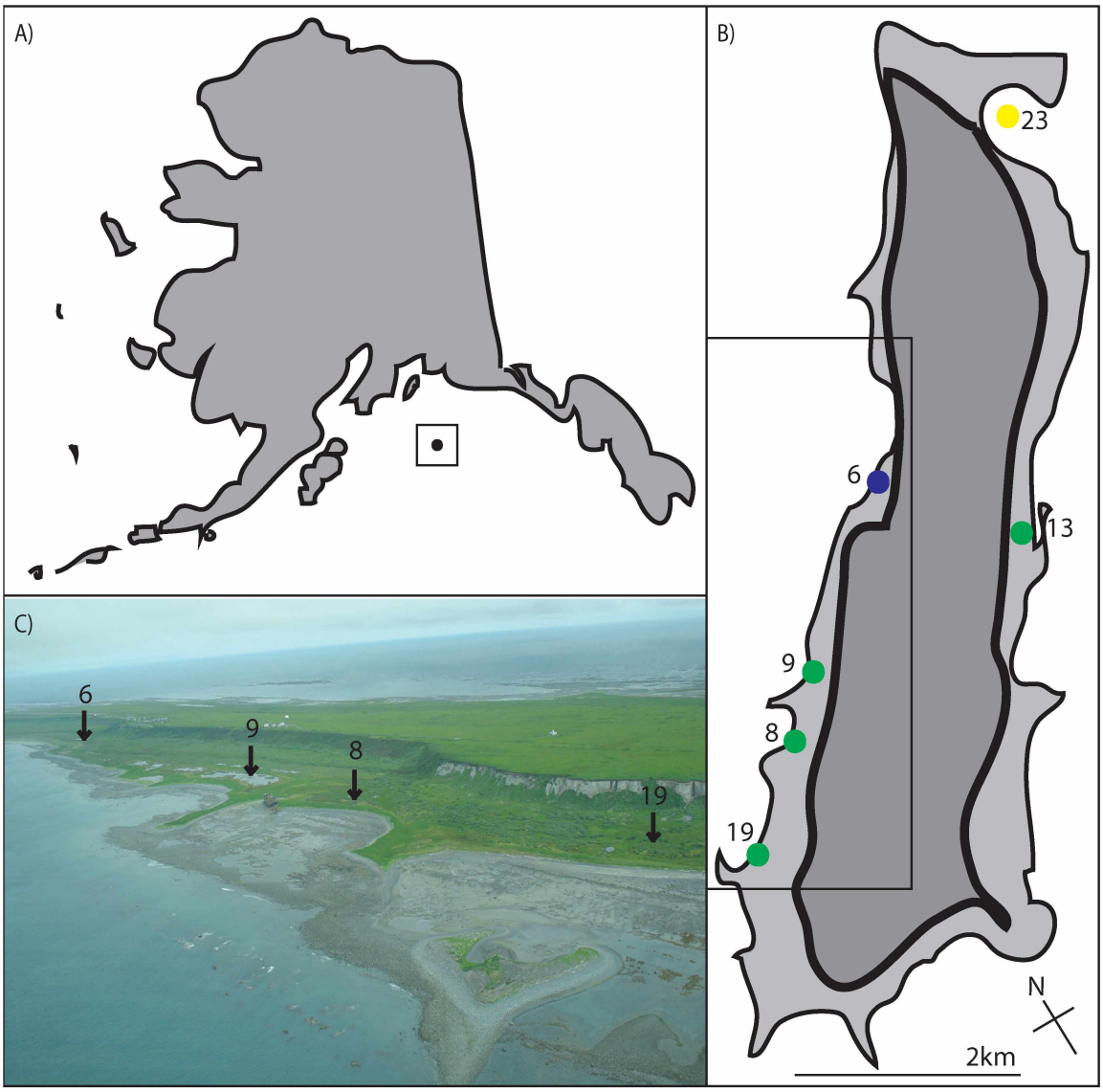

Figure 1.1. Collecting locations of threespine stickleback. A) Outline of Alaska with box highlighting the location of Middleton Island. B) Middleton Island in its current state (fine outline) with the approximate size of the island prior to the 1964 earthquake outlined inside (heavy line) and sites labeled. The yellow dot indicates the oceanic site, the blue dot indicates the freshwater site, and the green dots indicate sites containing individuals with both oceanic and freshwater phenotypes. North arrow marks true north. C) Photograph of the region of the island within the box in B, demonstrating uplift and locations of several sampling sites.

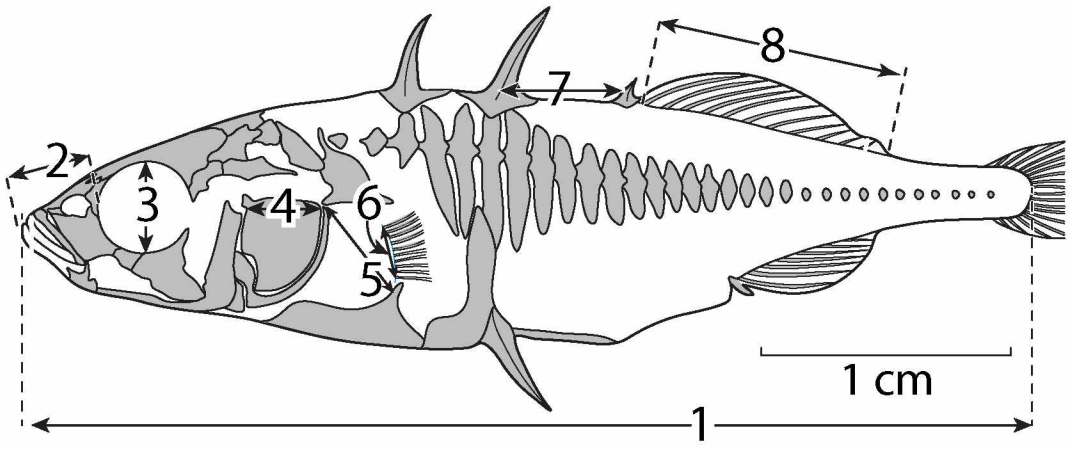

Figure 1.2. A suite of morphological and meristic traits associated with armor, feeding, and locomotion were analyzed. Digitally measured morphological traits: (1) Standard length, (2) Snout length, (3) Eye diameter, (4) Operculum width, (5) Diagonal length of pectoral fin abductor, (6) Length of the pectoral fin insertion, (7) Distance between second and third dorsal spines, (8) Length of the dorsal fin insertion.

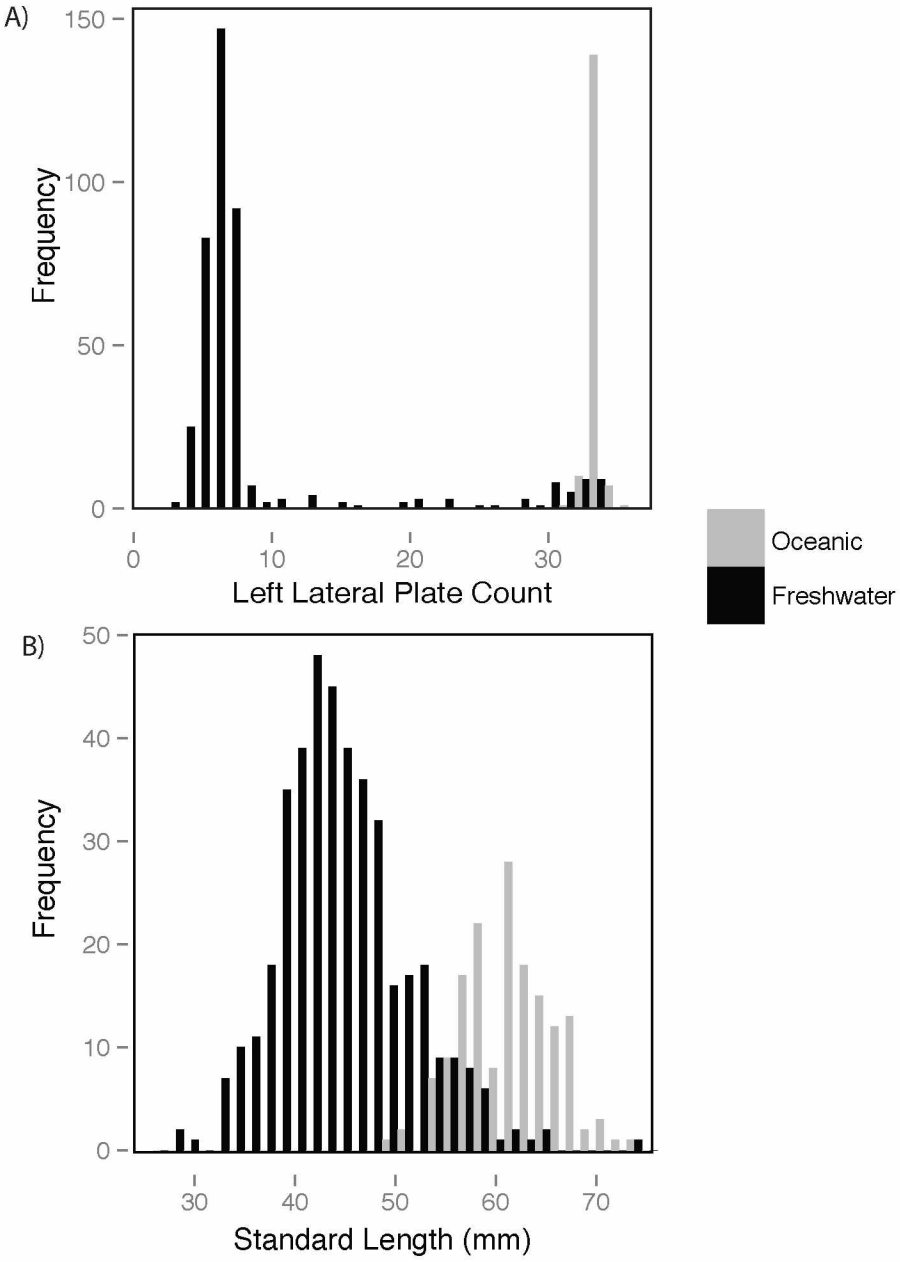

Figure 1.3. Freshwater and oceanic stickleback significantly differ in standard length and left lateral plate number. Frequency distributions of standard length (A) and left lateral plate number (B) across all sites pooled.

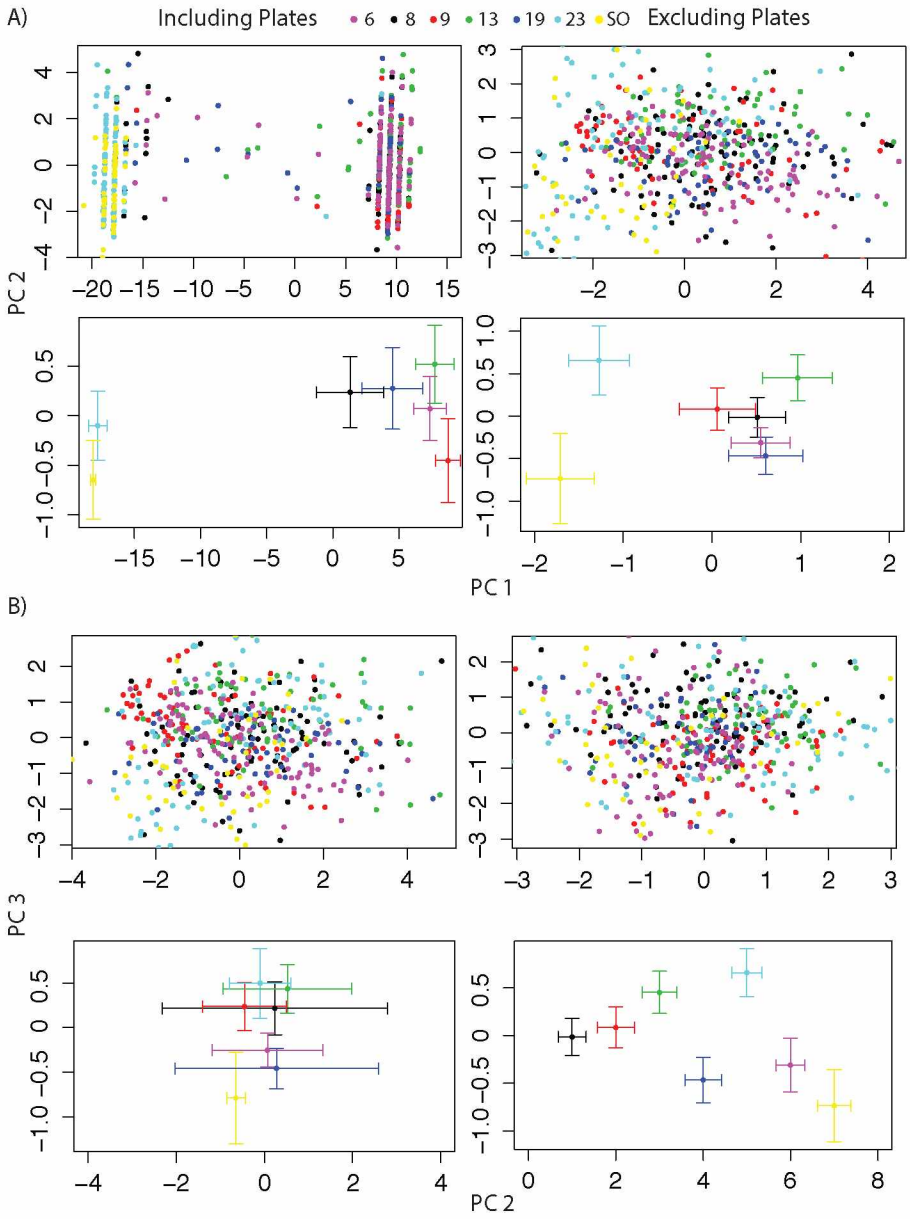

Figure 1.4. Principal components analyses (PCAs) were used to describe the overall distribution of phenotypic variation. Individuals are color-coded by site origin. In the upper panels, each dot represents an individual. In the lower panels, each dot represents the population mean +/- 2 SE. SO=oceanic individuals from sympatric sites. A) Principal components (PCs) 1 and 2. B) PCs 2 and 3.

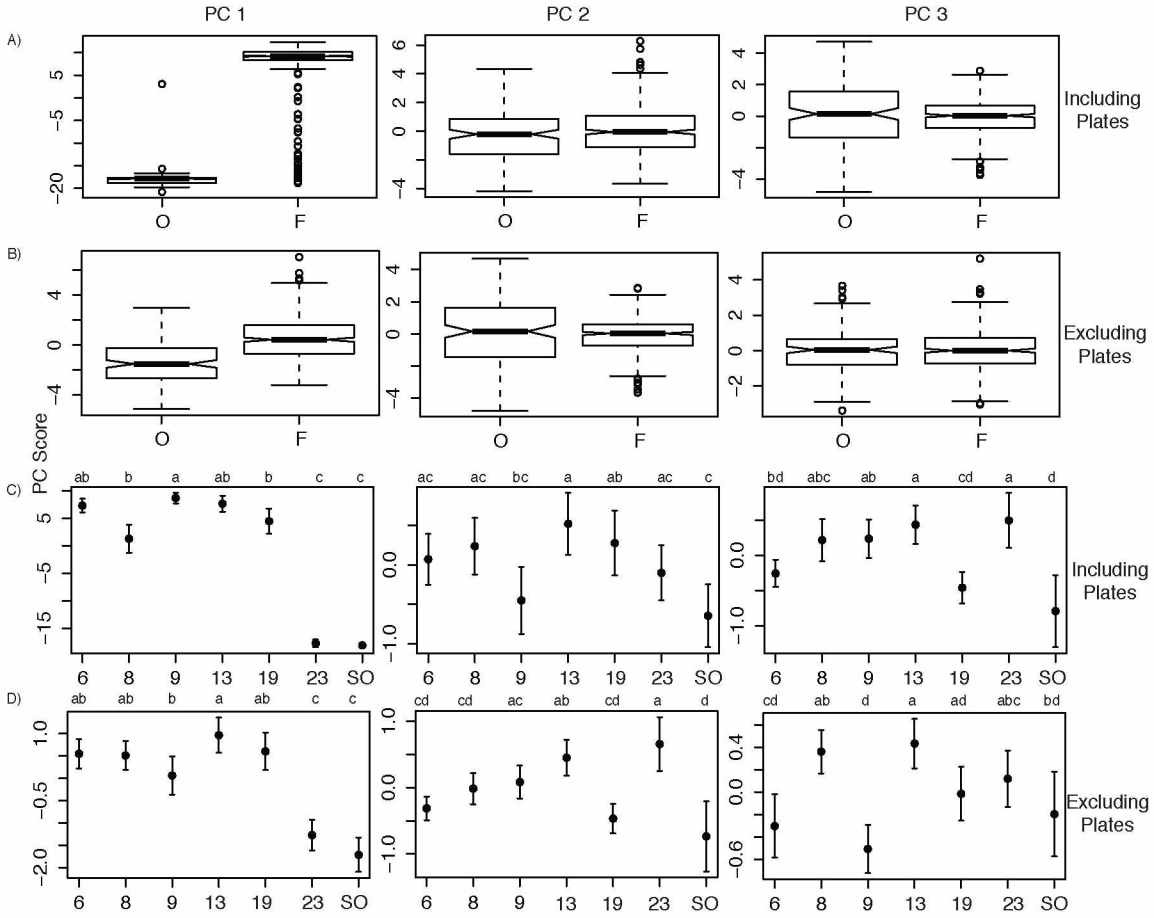

Figure 1.5. Ecotypes and populations significantly differ in mean principal component (PC) scores. A) and B) are boxplots of PC scores for ecotypes in the analyses including and excluding lateral plates. C) and D) are mean PC scores +/- 2 SE for populations in the analyses including and excluding lateral plates. Letters represent populations that significantly differ from each other based on Tukey's post hoc tests.

Table 1.1. Trait means by population. All measurements are in mm. FW = resident freshwater; O = oceanic.

Site	6	8		9		13		19		23
Variable	FW	FW	O	FW	O	FW	O	FW	O	O
SL	44.6	44.7	57.9	41.1	63.4	45.6	63.2	46.1	58.5	61.9
Pelvic Spine Length (left)	4.4	4.5	7.4	3.9	7.2	4.2	7.9	4.5	7.4	7.8
Second Dorsal Spine Length	3.2	3.1	5.0	2.9	5.1	3.1	5.7	3.2	5.2	5.5
Snout Length	3.1	3.3	4.2	3.0	4.8	3.3	5.2	3.3	4.8	4.8
Eye Diameter (left)	3.9	3.9	4.9	3.5	5.2	4.0	5.4	4.1	5.1	5.2
Operculum Width (left)	3.9	3.9	4.9	3.6	6.5	3.8	5.0	4.1	4.9	5.0
Diagonal of Pectoral Fin Muscle (left)	4.2	4.1	6.8	3.7	7.0	4.2	7.4	4.3	6.8	7.2
Pectoral Fin Insertion Length (left)	2.1	2.1	3.1	1.9	3.5	2.1	3.4	2.1	3.1	3.3
Length of First Lateral Plate behind Second Dorsal Spine (left)	4.4	5.2	7.4	4.7	7.9	4.8	8.0	5.0	7.3	7.5
Distance between 2nd-3rd Dorsal Spines	5.7	5.9	7.6	5.2	8.1	6.1	8.6	5.9	7.6	8.8
Dorsal Fin Length	8.9	9.3	12.7	8.9	13.5	9.3	13.8	9.6	12.8	13.8
Lateral Plate Number (left)	8.0	10.0	33.6	6.5	32.5	7.3	33.3	10.8	33.4	33.8
Number of Gill Rakers (right)	18.9	18.1	21.3	18.5	22.0	18.4	22.2	19.0	22.3	21.6
Number of Pectoral Fin Rays (left)	10.0	10.0	9.8	10.0	10.0	9.9	10.0	10.0	10.0	10.0
Number of Dorsal Fin Rays	10.3	10.3	11.3	10.7	11.5	10.3	11.6	10.6	11.7	11.5
Number of Anal Fin Rays	7.5	7.6	8.0	7.9	8.5	7.3	8.5	7.7	8.7	8.1
Number of Caudal Fin Rays	12.0	12.0	11.7	12.0	12.0	11.7	11.0	11.9	11.5	11.1
n	100	95	5	75	2	73	27	64	19	100

Table 1.2. Results of Tukey's Post-Hoc tests comparing principal component (PC) scores in the analyses including and excluding lateral plates. SO signifies oceanic individuals from sympatric sites. Bold represents significance at $p<0.05$.

		Including Plates			Excluding Plates		
		PC1	PC2	PC3	PC1	PC2	PC3
Site 6	Site 8	0.265	1.00	0.847	1.00	0.731	**0.001**
	Site 9	0.815	0.385	0.219	0.472	0.489	0.905
	Site 13	1.00	0.569	**0.016**	0.663	**0.005**	**<0.001**
	Site 19	0.084	0.987	0.965	1.00	0.992	0.696
	Site 23	**<0.001**	0.991	**0.002**	**<0.001**	**<0.001**	0.123
	SO	**<0.001**	0.151	0.236	**<0.001**	0.540	0.998
Site 8	Site 9	**0.011**	0.179	0.922	**<0.001**	**0.037**	**<0.001**
	Site 13	0.183	0.846	0.370	0.574	0.298	1.00
	Site 19	0.991	1.00	0.357	1.00	0.383	0.396
	Site 23	**<0.001**	0.898	0.142	**<0.001**	**0.011**	0.759
	SO	**<0.001**	0.060	**0.013**	**<0.001**	**0.037**	0.072
Site 9	Site 13	0.964	**0.007**	0.973	**0.018**	0.659	**<0.001**
	Site 19	**0.003**	0.138	**0.044**	0.474	0.217	0.151
	Site 23	**<0.001**	0.814	0.874	**<0.001**	0.091	**0.007**
	SO	**<0.001**	0.995	**<0.001**	**<0.001**	**0.017**	**0.038**
Site 13	Site 19	0.056	0.978	**0.002**	0.864	**0.002**	0.245
	Site 23	**<0.001**	0.182	1.00	**<0.001**	0.957	0.549
	SO	**<0.001**	**0.002**	**<0.001**	**<0.001**	**<0.001**	0.741
Site 19	Site 23	**<0.001**	0.790	**<0.001**	**<0.001**	**<0.001**	0.991
	SO	**<0.001**	**0.047**	0.845	**<0.001**	0.942	0.979
Site 23	SO	1.00	0.465	**<0.001**	0.718	**<0.001**	0.670

Chapter 2 Genetic Structure of 50 Year Old Stickleback Populations on Alaskan Earthquake-Uplifted Islands[1]

Abstract

How rapidly can animal populations in the wild evolve when faced with sudden environmental shifts? Uplift during the 1964 Great Alaska Earthquake abruptly created freshwater ponds on multiple islands in Prince William Sound and the Gulf of Alaska. In the short time since the earthquake, the phenotypes of resident freshwater threespine stickleback fish on three of these islands have changed dramatically from their oceanic ancestors. To test the hypothesis that these freshwater populations were derived from oceanic ancestors less than 50 years ago, we generated a set of over 130,000 single nucleotide polymorphism genotypes from 1,057 individuals using restriction site-associated DNA sequencing (RAD-seq). Population genetic analyses of these data support the hypothesis of recent and repeated independent colonization of freshwater habitats by oceanic ancestors. We found evidence of recurrent gene flow between oceanic and freshwater ecotypes where they co-occur. Our data implicate natural selection in phenotypic diversification, and support the hypothesis that the metapopulation organization of this species helps maintain a large pool of genetic variation that selection can rapidly act upon when oceanic stickleback colonize freshwater environments. We found that the freshwater populations, despite their young age, have diverged phenotypically from oceanic ancestors to nearly the same extent as populations that were likely founded thousands of years ago. Our results support the hypothesis that most

[1] Lescak, E.A., Bassham, S., Catchen, J., Gelmond, O., von Hippel, F.A., Sherbick, M.L., and Cresko, W.A. Prepared for submission to The Proceedings of the National Academy of Sciences.

stickleback evolution in fresh water occurs within the first few decades after invasion of a novel environment.

Introduction

On March 27th 1964, the largest earthquake ever recorded in North America struck the south coast of Alaska (1-3). This catastrophe uplifted islands in Prince William Sound and the Gulf of Alaska in just a few minutes, creating ponds from formerly marine habitat and setting the stage for the diversification of threespine stickleback fish (*Gasterosteus aculeatus*). This seismic disturbance provided an excellent opportunity to address long-standing evolutionary questions regarding how often dramatic phenotypic shifts can happen over contemporary time scales in the wild (4-13). Despite examples of rapid divergence in wild populations, evolutionary rates may often be constrained by a suite of factors. For example, evolution in new habitats may be limited by mutation rates (14-19). Even when adaptation occurs from standing genetic variation, evolution via selection of numerous independent loci of small effect may be time consuming (20-27).

We know, however, that evolution *can* occur rapidly, particularly under artificial selection or in human-altered landscapes (28-30). Most plants and animals have been domesticated in the last 10,000 years (31), many dog breeds diversified within the past few hundred years (32), insecticide resistance in arthropods has evolved within decades (33, 34), and numerous antibiotic resistant bacteria have arisen in the last few years (28, 35). In addition, empirical studies in the wild - particularly during significant environmental changes - have demonstrated that strong selection and rapid evolution over decades may be more common than once thought (12, 36-41) and may be of clear importance for organismal responses to anthropogenic perturbations and climate change (42).

A rapid evolutionary response is expected when the intensity of directional selection is strong (13), a scenario likely to occur immediately after a habitat shift or environmental disturbance (43-45). However, because of previous technological limitations, few studies have been able to use genetic data to fully disentangle evolution from induced phenotypic plasticity without accompanying genetic change. The small numbers of markers previously available for most population genetic studies have not provided the necessary precision with which to study very recently diverged populations (but see (9, 41)). As a consequence, the frequency of contemporary evolution in the wild is still poorly determined, and its genetic and genomic bases remain unclear.

The advent of new DNA sequencing technologies (46, 47) is beginning to allow for the precise inference from genomic data of colonization history and evolutionary patterns (48-52). Research using threespine stickleback has been at the forefront of using these new tools (53-57) and is ideal for testing hypotheses about contemporary evolution. Postglacial adaptive radiations over the last 12,000-20,000 years in newly invaded freshwater habitats have spawned divergent phenotypes that demonstrate parallel phenotypic evolution (58), with underlying parallel genetic (59-61) and genomic (53, 54, 56) bases. An open question, however, is whether this parallel divergence in stickleback actually requires thousands of years, or whether it can occur in the wild over contemporary time scales, as is implied by phenotypic studies of a small number of recently formed artificial and wild stickleback populations (62-68). Also unknown is how often the countless populations of stickleback in geographically close ponds represent invasion followed by local dispersal or independent founding from the oceanic ancestors.

To address these questions, we examined populations from three islands (Middleton, Montague, and Danger) in Prince William Sound and the Gulf of Alaska that were likely founded after the 1964 earthquake (Fig. 2.1; Table 2.1). Middleton Island was uplifted 3.4 m, creating new freshwater habitat on a previously submarine platform (3). Similarly, Danger and Montague islands experienced uplift and creation of new ponds (1). Stickleback now can be found in many of the habitats produced by the earthquake (69). We produced and analyzed RAD-seq data (70-72) from 25,000 loci in 1,057 individuals collected from a total of 20 populations (17 freshwater and three oceanic) from all three islands and one population in Cook Inlet. Individuals from these populations exhibit phenotypic variation characteristic of resident freshwater and oceanic ecotypes from elsewhere in the species' distribution (73-82). Deep sequencing of our large sample size yielded a set of 130,000 single nucleotide polymorphisms (SNPs) and a total of 146 million genotypes. This genomic data set allowed us to ask whether phenotypic and genetic divergence in stickleback, thought to require thousands of years, can occur in decades. We also tested whether replicated colonization of newly formed freshwater habitats by oceanic stickleback ancestors occurred independently among islands and over small geographic distances within a single island.

Results

Genetic variation was partitioned primarily between oceanic and freshwater ecotypes. To infer the evolutionary histories of populations on all three islands, we generated RAD libraries for 1,057 individuals from 21 total populations (20 originating from the islands plus one from Cook Inlet). These libraries were single-end sequenced in 12 lanes of an Illumina HiSeq2500 (Table 2.2), resulting in a total of more than more than 1.4 billion 101

bp sequences that passed several stringent quality filters. Each individual was represented by an average of 1.25 million sequences of which approximately 1 million (84%) were aligned to the reference genome, and nearly all (99%) were used by *Stacks* for downstream analyses. Approximately 130,000 SNPs were identified, most of which were called in every individual, thus providing a powerful data set for studies of recent and rapid evolution (Table 2.2).

We performed principal components analyses (PCA) using SNP data from the 21 populations. Similar to our phenotypic results (Lescak et al. unpublished data), we found that the major axis of genetic variation (PC1) represents a continuum of oceanic to freshwater genotypes and also separates freshwater populations from the three islands (Fig. 2.2-2.3). The optimum number of clusters (*K*) in the *STRUCTURE* analysis was 2, representing oceanic (red) and freshwater (blue) genotypes. Posterior probability of assignment to the oceanic genotype decreases, and probability of assignment to the freshwater genotype correspondingly increases, as populations increase in PC1 score.

Post-1964 populations were not derived from pre-existing freshwater populations. Historical maps and aerial photos show that no freshwater habitat existed before the 1964 earthquake on Danger Island. However, Middleton Island and the study area of Montague Island had pre-existing freshwater sites. Extensive sampling of Stump Lake on Montague Island yielded no fish of any species. However, stickleback were found in a pre-existing pond on Middleton (Mi12, Fig. 2.1), which rests on an upper terrace that formed approximately 2,400 years ago (3). Stickleback from this site were separated from the remaining sites in the study system along PC2 (Fig. 2.4-2.5), formed a clearly unique genotypic cluster in *STRUCTURE* (Fig. 2.3), and were the most differentiated freshwater

population on all three islands (Fig. 2.2-2.3, 2.5-2.6; Table 2.3-2.5). Site 12 is therefore unlikely to have founded the other freshwater populations on Middleton.

One hypothesis for the origin of freshwater stickleback in Mi12 is that they were unintentionally introduced during stocking of trout into this lake in the mid 20th century. The stocked trout originated from a hatchery on Upper Fire Lake (UFL) north of Anchorage, Alaska (Alaska Department of Fish and Game, personal communication). *STRUCTURE* analyses show that Mi12 and UFL represent distinct genotypic clusters with different allele frequencies (Fig. 2.7, Table 2.6), and the two sites are also highly divergent (F_{ST} = 0.164; Table 2.3). Taken together, these results support the hypothesis that stickleback from UFL founded neither Mi12 nor any other ponds on Middleton Island.

Independent evolution occurred quickly among islands and among regions within Middleton and Montague Islands. Using an Analysis of Molecular Variance (AMOVA; 83, 84) framework, we found that across all populations, most of the genetic variation (~74%) was partitioned among individuals, while 15% of the genetic variation was attributed to differences among sites within islands (Table 2.7). Despite having distinct phenotypic differences (Lescak et al. unpublished data), we found that only 5% of the genetic variation was partitioned between oceanic and freshwater habitats, but 18% of the variation was partitioned among populations within habitat types (freshwater or oceanic). Because the PCA and *STRUCTURE* results indicated genetic homogeneity among oceanic sites (Fig. 2.2), this partitioning of genetic variation must be due to differences among freshwater sites.

Furthermore, *STRUCTURE*, K-means clustering, PCA, and pairwise F_{ST} (Figs. 2.2-2.3, 2.6; Table 2.3-2.5) comparisons consistently support the presence of three distinct clusters of freshwater sites on Middleton Island. One group (designated FW1) contains sites Mi07

and Mi15. A second (FW2) consists of sites Mi06, Mi11, Mi16, and Mi28, as well as freshwater phenotype individuals from sites Mi08, Mi13 and Mi22. Mi14 supports a mix of freshwater individuals from FW1 and FW2. A third freshwater grouping is Mi12 (Table 2.8). Within Middleton Island, 9% of the variation among sites is attributed to these different regions, while individuals within sites and sites within regions each account for 5-6% of the variation. F_{ST} comparisons among populations within FW1 and FW2 range from 0.002-0.047, while those between pairs of populations among FW1, FW2 and Mi12 range from 0.015-0.198 (Table 2.3). Despite the recent time frame and small geographic distances, stickleback appear to have invaded and evolved freshwater phenotypes at least three times independently on an island only 8 km long and 1.6 km wide.

Similar to the results on Middleton, *STRUCTURE*, K means clustering, F_{ST} comparisons, and PCA support differentiation between Montague Island sites from the southeast (Mo35, Mo36, and Mo37; designated 'MoSE') and those from the southwest (Mo30 and Mo33, designated 'MoSW'; Fig. 2.3; Tables 2.3, 2.8, 2.9). The two populations that comprise MoSW were further separated from each other in *STRUCTURE* and K means analyses, and Mo33 appeared to consist of two freshwater populations (Figs. 2.3, 2.7; Table 2.9). Assignment to either MoSE or MoSW accounted for 17% of the variation among populations on Montague Island, while individuals within sites, and sites within MoSE and MoSW, accounted for approximately 3% and 5% of the variation, respectively. The remaining 75% of the variation is at the individual level (Table 2.7). Pairwise F_{ST} between populations within MoSE and MoSW ranges from 0.003-0.030, while F_{ST} between pairs of populations from MoSE and MoSW ranges from 0.037-0.125 (Table 2.3). The fine-scale population structure within Montague and Middleton Islands suggests that there were at

least five independent colonizations (Mi12, MiFW1, MiFW2, MoSE and MoSW) by oceanic ancestors on these two islands followed by rapid evolution. The presence of phenotypically oceanic (high plated) fish in a freshwater habitat (Da04) that form a unique genotypic cluster from nearby oceanic stickleback at Danger Island (Da02; Fig. 2.3; Table 2.10) argues that independent colonization of newly available habitats continues to occur.

Incongruent phenotypic and genetic variation indicated recurring introgressive hybridization. Phenotypically oceanic fish are sympatric with freshwater fish in five of the sites genotyped on Middleton Island (Fig. 2.1). Such ponds could house populations segregating phenotypes as polymorphisms, or could be zones of secondary contact between two genetically differentiated populations. *STRUCTURE* analyses revealed both oceanic and freshwater genotypes (Figs. 2.2-2.3), supporting the hypothesis of secondary contact of distinct populations. As expected, in these populations the genetic divergence between oceanic and freshwater ecotypes was highly correlated with lateral plate variation (Figs. 2.8-2.9), and a significant relationship exists between the genetic PC1 scores and lateral plate number ($r^2 = 0.54$, $t=-33.61$, $p<0.001$). Two main clusters of points comprised individuals with concordant freshwater or oceanic phenotypes and genotypes (Fig. 2.8). However, the presence of individuals in freshwater habitats with discordant phenotypes and genotypes indicates introgressive hybridization, trait plasticity, or individuals in the initial stages of divergence. In further support of introgressive hybridization, out of the 132 individuals surveyed from oceanic habitats on Danger (Da02) and Middleton Islands (Mi17 and Mi23), one individual had a freshwater genotype and another had a freshwater phenotype (Fig. 2.10). These anomalies could lend further credence to gene flow from

derived freshwater populations back into oceanic populations, or demonstrate the role of plasticity in phenotypic variation.

Discussion

Patterns of genetic differentiation supported independent evolution after the earthquake. Stickleback on three seismically uplifted islands harbor extensive genetic variation that was partitioned primarily among individuals. A significant amount of the remaining genetic variation, however, was also partitioned between fish with divergent lateral plate phenotypes in oceanic and freshwater habitats, as well as among those with freshwater phenotypes from freshwater sites (Table 2.7). Genetic divergence between oceanic and freshwater ecotypes has previously been documented across the stickleback holarctic distribution (54), as well as among populations in close proximity within Alaska (53, 85) and within Oregon (81). The overall divergence (as measured by F_{ST}) that we found here, however, is lower for several freshwater-oceanic pairs than has been documented previously, and is consistent with much more recent freshwater colonization from a panmictic oceanic ancestor (Fig. 2.11). In addition, genetic diversity was not much lower in freshwater habitats as compared to oceanic, indicating that either the initial colonizing populations were large, or (more likely) that recurrent gene flow has been occurring between oceanic and freshwater stickleback since the time the populations were founded (Table 2.11).

Our data did not support the possibility that the pre-1964 population, Mi12, was the progenitor of the younger freshwater populations on the island. In addition, we found no evidence that Mi12 or other populations on Middleton were colonized by stickleback from mainland UFL in the event stickleback were introduced inadvertently when trout were

stocked into Mi12. The two populations - UFL and Mi12 - were highly genetically divergent as well as phenotypically different (Fig. 2.7; Tables 2.3, 2.6). If individuals from UFL were introduced into a pre-existing stickleback population in Mi12, intra-population genetic structure would be expected, but we found none. Mi12 fish were polymorphic for lateral plate number, but *STRUCTURE* analysis of individuals from Mi12 revealed an optimum *K* of 1, and pairwise F_{ST} between low and completely plated individuals within the pond shows little genetic differentiation (F_{ST} = 0.004). Rather, the high genetic divergence between oceanic populations and Mi12 (F_{ST}~0.16; Table 2.3) was consistent with its founding at the time the terrace was formed by an older uplift event, approximately 2,400 years ago (3).

Our population genomic data clearly support a recent oceanic origin and subsequent differentiation of resident freshwater stickleback in post-1964 ponds on all three islands. In addition, intra-island genetic variation partitions freshwater sites from Middleton and Montague into three and two groups respectively (Mi12, MiFW1, MiFW2, and MoSE, MoSW) that are geographically distinct but, in the case of Middleton Island, still quite close. The two genetic groupings on Montague are found on opposite sides of the island (Fig. 2.1) and are separated by a mountain range, which likely prevents gene flow between them. However, gene flow may occur within watersheds, particularly the southeast watershed (Fig. 2.2-2.3; Tables 2.3, 2.4, 2.9). *STRUCTURE* analysis also revealed two distinct groups in Mo33 (Figs. 2.3, 2.9), likely representing two freshwater genotypes, since neither of these clusters is shared with oceanic individuals (Fig. 2.7). The southwest of Montague Island was uplifted an astounding 13-15 m as a result of the 1964 earthquake, which is much more displacement than Middleton Island experienced (3.4 m; (1)). Increased uplift may

have caused these Montague ponds to be more isolated from oceanic stickleback than the other post-1964 ponds, impeding secondary contact.

Overall, our findings support the hypothesis of at least six independent colonization events by oceanic ancestors within as well as among islands. Because we examined a subset of freshwater habitats on only three of the many islands and coastal regions throughout Prince William Sound and the Gulf of Alaska impacted by uplift, this is likely to be a significant underestimate of the number of times that freshwater stickleback have evolved independently in the last 50 years throughout this region. Furthermore, our results demonstrate a range of populations along the oceanic to freshwater phenotypic and genetic continua, from Da04, which closely resembles oceanic populations, to Mi07 and Mi15, which appear to have the most extreme freshwater phenotypes and genotypes (Figs. 2.2, 2.5). This spectrum likely reflects colonization and hybridization history. Da04 could be a more recently founded population that experiences regular gene flow with oceanic stickleback, while Mi07 and Mi15 could have been founded earlier and remained isolated.

Phenotypic differentiation supports a role for strong divergent selection. Despite being founded less than 50 years ago, freshwater stickleback populations on Middleton Island were differentiated from oceanic fish across multiple morphological features (Lescak et al. unpublished data). As expected, variation in lateral plate count was a major driver of the phenotypic divergence we measured. However, traits involved in foraging, defense and swimming also distinguished the two ecotypes (Table 1.1). Many of these traits have a known genetic basis (59-61, 77, 82, 86-89), with quantitative trait loci (QTL) mapping to multiple linkage groups (89-91), which supports that the phenotypic divergence we observed on Middleton Island is due at least in part to evolution.

Even in freshwater habitats where oceanic and freshwater ecotypes are sympatric, they were phenotypically divergent to nearly the same extent as allopatric ecotypes (Fig. 1.4, Table 1.4). Given the topography of Middleton Island and the locations of the newly formed freshwater habitats, secondary contact between oceanic and freshwater fish is the most probable cause of this sympatric co-occurrence. This scenario has likely been repeated over many years, and creates the potential for gene flow that would inhibit phenotypic divergence simply by drift (56, 92-96). The clearly discontinuous trait distributions between ecotypes are therefore most likely due to strong divergent selection on the phenotypes in the alternative oceanic and freshwater habitats, which is in agreement with other studies of selection in stickleback (76, 97, 98). The strength of selection (*s*) on the region containing *Eda* - the locus associated with lateral plate loss (60, 61, 74, 80, 89) - in oceanic stickleback transplanted to artificial freshwater ponds was found to be 0.52 (76). Even selection coefficients an order of magnitude smaller on other traits would maintain adaptive phenotypic differentiation in the face of all but the highest levels of gene flow.

Most stickleback phenotypic evolution may occur in the first decades after colonization. The post-glacial adaptive radiation of threespine stickleback in Cook Inlet is a well-described model of rapid evolution for which there is ample evidence of phenotypic and genetic divergence between ecotypes likely occurring over thousands of years (53, 61, 85, 91, 99-104). However, the independent colonization of newly available freshwater habitat by oceanic ancestors on Middleton, Montague, and Danger Islands has occurred on a time scale that is orders of magnitude shorter. While life history traits vary among stickleback populations, 50 years likely represents only 25 to a maximum of 50 generations

(99). Remarkably, during this short time span, freshwater and oceanic phenotypic divergence has reached nearly the same degree as measured in Cook Inlet populations that were probably founded about 13,000 years ago (61; Fig. 2.11). Our results demonstrate that evolution can occur in this species on decadal time scales, and pose the hypothesis that much of the previously documented post-glacial divergence of freshwater stickleback populations that has been inferred to have occurred over thousands of years (53, 73, 85, 91, 100, 103, 104) actually occurred during the first few generations following colonization (64, 67, 75, 101, 102).

While phenotypic differentiation in decades is a rate that is not often observed in nature (reviewed by 4, 6-8), it has previously been documented in threespine stickleback (64, 67-69, 101) as well as other fish species. For example, ancestral and derived populations of mosquitofish (*Gambusia affinis*) diverged in several life history traits in less than 60 years (105). Similarly, Pacific salmon (*Oncorhynchus* spp.) adapted to different breeding environments and developed partial reproductive isolation over a period of 13-26 generations (5, 7, 106). Parallel life history evolution has also been observed in the guppy (*Poecilia reticulata*) in only 4-11 years as a result of changes in predation pressure (107). Here, for the first time, we document the fine-scale population genetic patterns associated with rapid phenotypic evolution of stickleback after colonization of new habitats in the wild.

Introgressive hybridization maintains the pool of standing genetic variation. The possibility for recurrent gene flow between fish in different habitats persists. Our data document that oceanic fish are still entering freshwater ponds to breed or are becoming trapped due to high tides or storm surges, as has been observed in other regions of

southcentral Alaska (e.g., 75). Though most fish from freshwater sites containing both ecotypes were morphologically freshwater or oceanic, a few were intermediate in phenotype, and *STRUCTURE* results indicated a degree of hybridization in these ponds (Figs. 2.2-2.3, Table 2.4, 2.5). These results suggest that at least some individuals are likely early generation hybrids between the differentiated oceanic and freshwater ecotypes. Of the 132 individuals analyzed from the oceanic sites (Da02, Mi17 and Mi23), one fish from Danger Island had a freshwater lateral plate phenotype and one Middleton fish had a freshwater genome (Fig. 2.10), supporting gene flow between derived freshwater and ancestral oceanic populations. Thus, despite the clear phenotypic diversification caused by differential selection in the two habitats, gene flow via introgressive hybridization of stickleback in sympatry may decrease divergence in neutral genomic regions and help maintain a large pool of standing genetic variation available in oceanic fish for subsequent bouts of adaptation to new freshwater habitats.

Our data clearly support the 'transporter hypothesis' (108) that standing genetic variation in stickleback is maintained by low levels of recurrent hybridization between freshwater and oceanic stickleback. The loci of traits under selection map to different linkage groups (90), suggesting that numerous regions need to be reassembled. However, linkage disequilibrium created and maintained by strong selection in each habitat (85), perhaps abetted by structural variation in the stickleback genome, may significantly reduce the number of independent regions that need to be 'transported' and reassembled (109). Supporting this hypothesis is our finding that parallel evolution in this species happens not only on the scale of different continents or dispersed geographic regions as has been amply documented (53, 56, 60, 80, 88, 110-112), but also on such small spatial scales as nearby

islands or even geographically proximate ponds on the same island. These data argue that rapid parallel evolution over decades in stickleback may occur frequently because it is underlain by a pliant genomic architecture that is itself the product of millions of years of evolution. If the findings from stickleback are generalizable to other systems, then contemporary evolution in the wild may be more common than previously documented.

Materials and Methods

Site selection. By comparing pre- and post-1964 maps and aerial imagery (Aerometrics, USGS, BLM) of islands in Prince William Sound and the Gulf of Alaska, we identified ponds on terrain that had been submarine prior to the 1964 Great Alaska Earthquake. Multiple water bodies fitting this criterion were found on Montague and Danger Islands, in Prince William Sound, and on Middleton Island, in the Gulf of Alaska (Fig. 2.1; Table 2.1). Middleton Island also has a pre-1964 pond (Mi12) on an upper terrace formed approximately 2,400 years ago (3). Since this site was stocked with rainbow trout from the Upper Fire Lake (UFL) hatchery (Eagle River, Alaska) in the 1960s, we also included stickleback samples from UFL in our analysis.

Field collections. Collections from 21 sites were made in the summers of 2005 (Montague and Middleton), 2010 (Danger and Middleton) and 2011 (Middleton and UFL). Threespine stickleback were collected using 0.32 and 0.64 cm mesh minnow traps set near shore and left overnight, euthanized with an overdose of MS-222 anesthetic, and preserved in 95% ethanol. Salinity was measured at each site by YSI-80 meter and ranged from 0.1-1.4 ppt for freshwater sites and 21.4-26.4 ppt for oceanic sites. We sampled one oceanic and one post-1964 freshwater site on Danger Island, five post-1964 freshwater sites on

Montague Island, and two oceanic sites, eleven post-1964 freshwater sites, and one pre-1964 freshwater site on Middleton Island (Fig. 2.1, Table 2.1).

Sample preparation. Caudal and pectoral fins were clipped for DNA extraction using the Qiagen DNeasy kit. DNA and corresponding bodies were assigned a unique identification number for association of genotype with phenotype. Bodies were fixed in 10% neutral buffered formalin for at least 48 hours, bleached in a 0.05% hydrogen peroxide solution, stained in a 0.1% Alizarin red S solution, destained in 1% KOH, and preserved in 70% ethanol. Individuals were classified as having oceanic or freshwater phenotypes by a principal components analysis using R (113).

RAD library preparation and sequence analysis. Genomic DNA from each of the 1,057 individuals from all three islands and UFL was digested with the restriction enzyme SbfI-HF (NEB), and RAD-seq libraries were created as previously reported (53, 70, 71). Uniquely barcoded samples representing 76 to 96 individuals were run per lane in 12 total lanes of sequencing on an Illumina HiSeq2500 platform, and on average, each lane resulted in approximately 157 million SE sequences, of which about 113 million were retained (72%). The 101 nucleotide long reads included 6 nucleotide in-line barcodes to identify individual fish. Raw sequence data were demultiplexed by barcode and filtered for quality using the process_radtags program in the *Stacks* software suite (114, 115). Reads were aligned against the stickleback reference genome (version BROADs1, Ensembl release 64) using GSnap (116, 117), allowing for up to 5 mismatches and gaps of length 2, disabling terminal alignments, and requiring unique alignments. The alignments were processed and genotypes were called for each locus across all samples using the pstacks, cstacks, and sstacks programs from *Stacks*. For a locus to be included in further analyses, we required

that it be present in all populations and successfully genotyped in at least 75% of individuals from each population. See Table 2.2 for tallies of raw and retained reads.

Statistical approach. Population genetic statistics (major allele frequency, percent polymorphic loci, and Wright's F statistics F_{IS} and F_{ST}) were calculated for every single nucleotide polymorphism (SNP) using the populations program in *Stacks* (114, 115). For bi-allelic SNP markers, π is a measure of expected heterozygosity, which is an overall indication of a population's genetic diversity. F_{IS} measures the reduction in observed heterozygosity compared to that which is expected for a locus in a population and can indicate nonrandom mating or cryptic population structure (118, 119).

To analyze population structure, we used the populations program in *Stacks* to output filtered SNP data from all RAD loci across all populations into a file formatted for *STRUCTURE* (120-122). Because of computational limitations, we randomly chose three subsets of 1,000 SNPs each to complete the analysis. These three subsets were found to yield comparable results, so results from only one subset are included. *STRUCTURE* analyses were performed on all sites together, as well as separately on subsets of the data for fine-scale detection of population structure (Figs. 2.2-2.3, 2.7; Tables 2.4-2.6, 2.9, 2.10). For all analyses, 10,000 burn-in steps and 10,000 replicates were used, with 10 runs for each potential *K* (number of genotypic groups). The optimal *K* for each analysis was chosen using the deltaK method (123). This same set of SNPs was used in *GenoDive* (124) to conduct K means cluster analyses using the same ranges of *K* tested in *STRUCTURE*. We also tested a higher level of population structure using Analyses of Molecular Variance (AMOVA; 83, 84) that nested sites within islands, within habitat type (defined by oceanic versus

freshwater), or within mean lateral plate morph (complete, partial, or low). Lastly, we used PCA to identify the major axes of genetic variation.

Acknowledgements

We thank M.S. Christy, S.A. Hatch, B. Lohman, V.M. Padula, L. Smayda, and K. Walton for assistance with fieldwork logistics and fish collection, as well as T. Wilson and M. Currey for laboratory help. The Chugach National Forest, Bureau of Land Management, Federal Aviation Administration, Anchorage Museum of History and Art, and Army Corps of Engineers provided access to their imagery collections. We thank J.J. Colgren, M. Currey, B.R. Harrison, J.A. Lopez, R.B. Lucas, and M.V. McPhee, as well as other members of the Cresko, Phillips and Streisfeld Labs at the University of Oregon, for discussions and comments. Fish were collected under Alaska Department of Fish and Game permit numbers SF-2005-020, SF-2010-029, and SF-2011-153 and all research protocols with vertebrates were approved by the University of Alaska Anchorage and University of Oregon Institutional Animal Care and Use Committees. This research was supported primarily by National Science Foundation DEB 0949053 to WAC and DEB 0919234 to FAvH, as well as IOS 102728 (WAC). Additional support came from a University of Alaska Anchorage Faculty Development Grant (FAvH), NIH grant 1R24GM079486-01A1 (WAC), NIH NRSA Ruth L. Kirschstein fellowship F32GM095213-01 (JMC), and the M. J. Murdock Charitable Trust (WAC). EAL was supported with funds from the University of Alaska Center for Global Change and Arctic Systems Research, the University of Alaska Anchorage, LGL Limited, the Rasmuson Fisheries Board, the American Fisheries Society, and the NSF Alaska EPSCoR Landscape Genetics program.

References

1. Plafker G (1967) *Surface faults on Montague Island associated with the 1964 Alaska earthquake* (US Government Printing Office).
2. Plafker G (1969) The Alaska Earthquake, March 27, 1964: regional effects – the tectonics of the March 27, 1964 Alaska Earthquake. 543-I:1–74.
3. Plafker G, Rubin M (1978) Uplift history and earthquake recurrence as deduced from marine terraces on Middleton Island, Alaska. 78:687–721.
4. Hendry AP, Kinnison MT (1999) The pace of modern life: measuring rates of contemporary microevolution. *Evolution* 53:1637–1653.
5. Hendry AP, Wenburg JK, Bentzen P, Volk EC, Quinn TP (2000) Rapid evolution of reproductive isolation in the wild: evidence from introduced salmon. *Science* 290:516–519.
6. Hendry AP, Kinnison MT (2001) An introduction to microevolution: rate, pattern, process. *Genetica* 112-113:1–8.
7. Kinnison MT, Hendry AP (2001) The pace of modern life II: from rates of contemporary microevolution to pattern and process. *Genetica* 112-113:145–164.
8. Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. *Genetica* 112-113:183–198.
9. Koskinen MT, Haugen TO, Primmer CR (2002) Contemporary fisherian life-history evolution in small salmonid populations. *Nature* 419:826–830.

10. Latta RG (2014) Using QTL and GBS SNPs to Study the Contemporary Evolution of Introduced Wild Oats in California. Presented at Plant and Animal Genome XXII Meeting San Diego, CA.
11. Thomann M, Imbert E, Engstrand RC, Cheptou P-O (2015) Contemporary evolution of plant reproductive strategies under global change is revealed by stored seeds. *J Evol Biol* 766–778.
12. Thompson JN (1998) Rapid evolution as an ecological process. *Trends Ecol Evol* 13:329–332.
13. Hairston Jr NG, Ellner SP, Geber MA, Yoshida T, Fox JA (2005) Rapid evolution and the convergence of ecological and evolutionary time. *Ecol Lett* 8:1114–1127.
14. Burke MK (2012) How does adaptation sweep through the genome? Insights from long-term selection experiments. *Proc R Soc B* 279:5029–5038.
15. Wright S (1931) Evolution in Mendelian populations. *Genetics* 16:97–159.
16. Kimura M, Ota T (1971) Theoretical aspects of population genetics. *Monogr Popul Biol* 4:1–219.
17. Nei M (1973) Analysis of gene diversity in subdivided populations. *Proc Natl Acad Sci USA* 70:3321–3323.
18. Radwan J, Babik W (2012) The genomics of adaptation. *Proc R Soc B* 279:5024–5028.
19. Aylmer FR (1930) The genetical theory of natural selection.
20. Hermisson J, Pennings PS (2005) Soft sweeps: molecular population genetics of adaptation from standing genetic variation. *Genetics* 169:2335–2352.
21. Pennings PS, Hermisson J (2006) Soft sweeps II--molecular population genetics of adaptation from recurrent mutation or migration. *Mol Biol Evol* 23:1076–1084.

22. Pennings PS, Hermisson J (2006) Soft sweeps III: the signature of positive selection from recurrent mutation. *PLoS Genetics* 2:e186–e186.

23. Barrett RDH, Schluter D (2008) Adaptation from standing genetic variation. *Trends Ecol Evol* 23:38–44.

24. Messer PW, Petrov DA (2013) Population genomics of rapid adaptation by soft selective sweeps. *Trends Ecol Evol* 28:659–669.

25. Mather K (1943) Polygenic inheritance and natural selection. *Biol Rev* 18:32–64.

26. Thoday JM, Thompson Jr JN (1976) The number of segregating genes implied by continuous variation. *Genetica* 46:335–344.

27. Mackay TFC, Stone EA, Ayroles JF (2009) The genetics of quantitative traits: challenges and prospects. *Nat Rev Genet* 10:565–577.

28. Lenski RE, Travisano M (1994) Dynamics of adaptation and diversification: a 10,000-generation experiment with bacterial populations. *Proc Natl Acad Sci USA* 91:6808–6814.

29. Dettman JR, Rodrigue N, Melnyk AH, Wong A, Bailey SF, Kassen R (2012) Evolutionary insight from whole-genome sequencing of experimentally evolved microbes. *Mol Ecol* 21:2058–2077.

30. Kohn LM, Anderson JB (2014) The underlying structure of adaptation under strong selection in 12 experimental yeast populations. *Eukaryot Cell* 13:1200–1206.

31. Palumbi SR (2001) Humans as the world's greatest evolutionary force. *Science* 293:1786–1790.

32. Boyko AR (2011) The domestic dog: man's best friend in the genomic era. *Genome Biol* 12:216.

33. Mallet J (1989) The evolution of insecticide resistance: have the insects won? *Trends Ecol Evol* 4:336–340.

34. Ffrench-Constant RH, Daborn PJ, Le Goff G (2004) The genetics and genomics of insecticide resistance. *Trends in Genetics : TIG* 20:163–170.

35. Graves JL, Tajkarimi M, Cunningham Q, Campbell A, Nonga H, Harrison SH, Barrick JE (2015) Rapid evolution of silver nanoparticle resistance in *Escherichia coli*. *Frontiers in Genetics* 6:42–42.

36. Roels SAB, Kelly JK (2011) Rapid evolution caused by pollinator loss in *Mimulus guttatus*. *Evolution* 65:2541–2552.

37. Jalvingh KM, Chang PL, Nuzhdin SV, Wertheim B (2014) Genomic changes under rapid evolution: selection for parasitoid resistance. *Proc R Soc B* 281:20132303–20132303.

38. Stuart YE, Campbell TS, Hohenlohe PA, Reynolds RG, Revell LJ, Losos JB (2014) Rapid evolution of a native species following invasion by a congener. *Science* 346:463–466.

39. Huey RB, Gilchrist GW, Carlson ML, Berrigan D, Serra L (2000) Rapid evolution of a geographic cline in size in an introduced fly. *Science* 287:308–309.

40. O'Steen S, Cullum AJ, Bennett AF (2002) Rapid evolution of escape ability in Trinidadian guppies (*Poecilia reticulata*). *Evolution* 56:776–784.

41. Lee CE (1999) Rapid and repeated invasions of fresh water by the copepod *Eurytemora affinis*. *Evolution* 53: 1423-1434.

42. Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008) Climate change and evolution: disentangling environmental and genetic responses. *Mol Ecol* 17:167–178.

43. Turcotte MM, Reznick DN, Hare JD (2011) The impact of rapid evolution on population dynamics in the wild: experimental test of eco-evolutionary dynamics. *Ecol Lett* 14:1084–1092.

44. Wendling CC, Wegner KM (2015) Adaptation to enemy shifts: rapid resistance evolution to local Vibrio spp. in invasive Pacific oysters. *Proc R Soc B* DOI: 10.1098/rspb.2014.2244.

45. Losas JB, Warheitf KI, Schoener TW (1997) Adaptive differentiation following experimental island colonization in Anolis lizards. *Nature* 387: 70-73.

46. Shendure J, Ji H (2008) Next-generation DNA sequencing. *Nat Biotechnol* 26:1135–1145.

47. Mardis ER (2008) Next-generation DNA sequencing methods. *Annu Rev Genomics Hum Genet* 9:387–402.

48. Hohenlohe PA, Phillips PC, Cresko WA (2010) Using population genomics to detect selection in natural populations: key concepts and methodological considerations. *Int J Plant Sci* 171:1059–1071.

49. Davey JW, Hohenlohe PA, Etter PD, Boone JQ, Catchen JM, Blaxter ML (2011) Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat Rev Genet* 12:499–510.

50. Fumagalli M, Vieira FG, Korneliussen TS, Linderoth T, Huerta-Sanchez E, Albrechtsen A, Nielsen R (2013) Quantifying population genetic differentiation from next-generation sequencing data. *Genetics* 195:979–992.

51. Lemmon AR, Lemmon EM (2012) High-throughput identification of informative nuclear loci for shallow-scale phylogenetics and phylogeography. *Syst Biol* 61:745–761.

52. Stapley J, Reger J, Feulner PGD, Smadja C, Galindo J, Ekblom R, Bennison C, Ball AD, Beckerman AP, Slate J (2010) Adaptation genomics: the next generation. *Trends in ecology & evolution (Personal edition)* 25:705–712.

53. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA (2010) Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* 6:e1000862–e1000862.

54. Jones FC, Grabherr MG, Chan YF, Russell P, Mauceli E, Johnson J, Swofford R, Pirun M, Zody MC, White S *et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* 484:55–61.

55. Arnegard ME, McGee MD, Matthews B, Marchinko KB, Conte GL, Kabir S, Bedford N, Bergek S, Chan YF, Jones FC *et al.* (2014) Genetics of ecological divergence during speciation. *Nature* 511:307–311.

56. Roesti M, Gavrilets S, Hendry AP, Salzburger W, Berner D (2014) The genomic signature of parallel adaptation from shared genetic variation. *Mol Ecol* 23:3944–3956.

57. Terekhanova NV, Logacheva MD, Penin AA, Neretina TV, Barmintseva AE, Bazykin GA, Kondrashov AS, Mugue NS (2014) Fast evolution from precast bricks: genomics of young freshwater populations of threespine stickleback Gasterosteus aculeatus. *PLoS Genetics* 10:e1004696.

58. Cresko WA, Mcguigan KL, Phillips PC, Postlethwait JH (2007) Studies of threespine stickleback developmental evolution: progress and promise. *Genetica* 129:105–126.

59. Chan YF, Marks ME, Jones FC, Villarreal G, Shapiro MD, Brady SD, Southwick AM, Absher DM, Grimwood J, Schmutz J *et al.* (2010) Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. *Science (New York, NY)* 327:302–305.

60. Colosimo PF, Hosemann KE, Balabhadra S, Villarreal G, Dickson M, Grimwood J, Schmutz J, Myers RM, Schluter D, Kingsley DM (2005) Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science (New York, NY)* 307:1928–1933.

61. Cresko WA, Amores A, Wilson C, Murphy J, Currey M, Phillips P, Bell MA, Kimmel CB, Postlethwait JH (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc Natl Acad Sci USA* 101:6050–6055.

62. Klepaker T (1993) Morphological changes in a marine population of threespined stickleback, *Gasterosteus aculeatus*, recently isolated in fresh water. *Can J Zool* 71: 1251-1258.

63. Kristjansson BK (2005) Rapid morphological changes in threespine stickleback, *Gasterosteus aculeatus*, in freshwater. *Environ Biol Fish* 74:357–363.

64. Bell MA, Aguirre WE, Buck NJ (2004) Twelve years of contemporary armor evolution in a threespine stickleback population. *Evolution; international journal of organic evolution* 58:814–824.

65. Kitano J, Bolnick DI, Beauchamp DA, Mazur MM, Mori S, Nakano T, Peichel CL (2008) Reverse evolution of armor plates in the threespine stickleback. *Curr Biol* 18:769–774.

66. Barrett RDH, Paccard A, Healy TM, Bergek S, Schulte PM, Schluter D, Rogers SM (2011) Rapid evolution of cold tolerance in stickleback. *Proc R Soc B* 278:233–238.

67. Aguirre WE, Bell MA (2012) Twenty years of body shape evolution in a threespine stickleback population adapting to a lake environment. *Biol J Linn Soc* 105: 817-831.

68. Bell MA, Aguirre WE (2013) Contemporary evolution, allelic recycling, and adaptive radiation of the threespine stickleback. *Evol Ecol Res* 15: 377-411.

69. Gelmond O, von Hippel FA, Christy MS (2009) Rapid ecological speciation in three-spined stickleback Gasterosteus aculeatus from Middleton Island, Alaska: the roles of selection and geographic isolation. *J Fish Biol* 75:2037–2051.

70. Baird NA, Etter PD, Atwood TS, Currey MC, Shiver AL, Lewis ZA, Selker EU, Cresko WA, Johnson EA (2008) Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS ONE* 3:e3376–e3376.

71. Etter PD, Bassham S, Hohenlohe PA, Johnson EA, Cresko WA (2011) SNP discovery and genotyping for evolutionary genetics using RAD sequencing. *Methods Mol Biol* 772:157–178.

72. Rowe HC, Renaut S, Guggisberg A (2011) RAD in the realm of next-generation sequencing technologies. *Mol Ecol* 20:3499–3502.

73. Bell MA, Foster SA (1994) *The Biology of the Stickleback.* Oxford University Press.

74. Bell MA (2001) Lateral plate evolution in the threespine stickleback: getting nowhere fast. *Genetica* 112-113:445–461.

75. von Hippel FA, Weigner H (2004) Sympatric anadromous-resident pairs of threespine stickleback species in young lakes and streams at Bering Glacier, Alaska. *Behaviour* 141: 1441-1464.

76. Barrett RDH, Rogers SM, Schluter D (2008) Natural selection on a major armor gene in threespine stickleback. *Science* 322:255–257.

77. Miller CT, Glazer AM, Summers BR, Blackman BK, Norman AR, Shapiro MD, Cole BL, Peichel CL, Schluter D, Kingsley DM (2014) Modular skeletal evolution in sticklebacks is controlled by additive and clustered quantitative trait Loci. *Genetics* 197:405–420.

78. McKinnon JS, Rundle HD (2002) Speciation in nature: the threespine stickleback model systems. *Trends Ecol Evol* 17:480–488.

79. Hagen DW (1967) Isolating mechanisms in threespine sticklebacks (*Gasterosteus*). *J Fish Board Can* 24:1637–1692.

80. Colosimo PF, Peichel CL, Nereng K, Blackman BK, Shapiro MD, Schluter D, Kingsley DM (2004) The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology* 2:E109.

81. Catchen J, Bassham S, Wilson T, Currey M, O'Brien C, Yeates Q, Cresko WA (2013) The population structure and recent colonization history of Oregon threespine stickleback determined using restriction-site associated DNA-sequencing. *Mol Ecol* 22:2864–2883.

82. McGuigan K, Nishimura N, Currey M, Hurwit D, Cresko WA (2011) Cryptic genetic variation and body size evolution in threespine stickleback. *Evolution* 65:1203–1211.

83. Excoffier L, Smouse PE, Quattro JM (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics* 131:479–491.

84. Excoffier L, Lischer HEL (2010) Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Molecular ecology resources* 10:564–567.

85. Hohenlohe PA, Bassham S, Currey M, Cresko WA (2012) Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback genomes. *Phil Trans R Soc B* 367:395–408.

86. Shapiro MD, Marks ME, Peichel CL, Blackman BK, Nereng KS, Jónsson B, Schluter D, Kingsley DM (2004) Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* 428:717–723.

87. Miller CT, Beleza S, Pollen AA, Schluter D, Kittles RA, Shriver MD, Kingsley DM (2007) cis-Regulatory changes in Kit ligand expression and parallel evolution of pigmentation in sticklebacks and humans. *Cell* 131:1179–1189.

88. Glazer AM, Cleves PA, Erickson PA, Lam AY, Miller CT (2014) Parallel developmental genetic features underlie stickleback gill raker evolution. *EvoDevo* 5:19.

89. Mills MG, Greenwood AK, Peichel CL (2014) Pleiotropic effects of a single gene on skeletal development and sensory system patterning in sticklebacks. *EvoDevo* 5:5.

90. Liu J, Shikano T, Leinonen T, Cano JM, Li M-H, Merilä J (2014) Identification of major and minor QTL for ecologically important morphological traits in three-spined sticklebacks (*Gasterosteus aculeatus*). *G3* 4:595–604.

91. Kimmel CB, Hohenlohe PA, Ullmann B, Currey M, Cresko WA (2012) Developmental dissociation in morphological evolution of the stickleback opercle. *Evol Devel* 14:326–337.

92. Nosil P, Funk DJ, Ortiz-Barrientos D (2009) Divergent selection and heterogeneous genomic divergence. *Mol Ecol* 18:375–402.

93. Turner TL, Hahn MW (2010) Genomic islands of speciation or genomic islands and speciation? *Mol Ecol* 19:848–850.

94. Cruickshank TE, Hahn MW (2014) Reanalysis suggests that genomic islands of speciation are due to reduced diversity, not reduced gene flow. *Mol Ecol* 23:3133–3157.

95. Via S, West J (2008) The genetic mosaic suggests a new role for hitchhiking in ecological speciation. *Mol Ecol* 17:4334–4345.

96. Via S (2009) Natural selection in action during speciation. *Proc Natl Acad Sci* 106 Suppl 1:9939–9946.

97. Schluter D, Marchinko KB, Barrett RDH, Rogers SM (2010) Natural selection and the genetics of adaptation in threespine stickleback. *Phil Trans R Soc B* 365:2479–2486.

98. Rennison DJ, Heilbron K, Barrett RDH, Schluter D (2015) Discriminating selection on lateral plate phenotype and its underlying gene, Ectodysplasin, in threespine stickleback. *Am Nat* 185:150–156.

99. Baker JA, Foster SA, Heins DC, Bell MA, King RW (1998) Variation in female life-history traits among Alaskan populations of the threespine stickleback, *Gasterosteus aculeatus* L. (Pisces: Gasterosteidae). *Biol J Linn Soc* 63:141–159.

100. Bell MA, Francis RC, Havens AC (1985) Pelvic reduction and its directional asymmetry in threespine sticklebacks from the Cook Inlet Region, Alaska. *Copeia* 1985:437–444.

101. Furin CG, von Hippel FA, Bell MA (2012) Partial reproductive isolation of a recently derived resident-freshwater population of threespine stickleback (*Gasterosteus aculeatus*) from its putative anadromous ancestor. *Evolution* 66:3277–3286.

102. Karve AD, von Hippel FA, Bell MA (2008) Isolation between sympatric anadromous and resident threespine stickleback species in Mud Lake, Alaska. *Environ Biol Fish* 81:287–296.

103. Jones FC, Brown C, Pemberton JM, Braithwaite VA (2006) Reproductive isolation in a threespine stickleback hybrid zone. *J Evol Biol* 19:1531–1544.

104. Aguirre WE, Ellis KE, Kusenda M (2008) Phenotypic variation and sexual dimorphism in anadromous threespine stickleback: implications for postglacial adaptive radiation. *Biol J Linn Soc* 95: 465-478.

105. Stockwell CA, Weeks SC (1999) Translocations and rapid evolutionary responses in recently established populations of western mosquitofish (*Gambusia affinis*). *Anim Conserv* 2:103–110.

106. Quinn TP, Hendry AP, Buck GB (2001) Balancing natural and sexual selection in sockeye salmon: interactions between body size, reproductive opportunity and vulnerability to predation by bears. *Evol Ecol R* 3: 917-937.

107. Reznick DN, Shaw FH, Rodd FH, Shaw RG (1997) Evaluation of the rate of evolution in natural populations of guppies (*Poecilia reticulata*) *Science* 275: 1934-1937.

108. Schluter D, Conte GL (2009) Genetics and ecological speciation. *Proc Natl Acad Sci* 106:9955–9962.

109. Feulner PG, Chain FJ, Panchal M, Elzaguirre C, Kalbe M, Lenz TL, Mundry M, Samonte IE, Stoll M, Milinski M, Reusch TB, Bomberg-Bauer E (2013) *Mol Ecol* 22: 635-649.

110. Rundle HD, Nagel L, Wenrick Boughman J, Schluter D (2000) Natural selection and parallel speciation in sympatric sticklebacks. *Science* 287:306–308.

111. Boughman JW, Rundle HD, Schluter D (2005) Parallel evolution of sexual isolation in sticklebacks. *Evolution* 59:361–373.

112. Kaeuffer R, Peichel CL, Bolnick DI, Hendry AP (2012) Parallel and nonparallel aspects of ecological, phenotypic, and genetic divergence across replicate population pairs of lake and stream stickleback. *Evolution* 66:402–418.

113. R Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing.

114. Catchen JM, Amores A, Hohenlohe P, Cresko W, Postlethwait JH (2011) Stacks: building and genotyping Loci de novo from short-read sequences. *G3* 1:171–182.

115. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA (2013) Stacks: an analysis tool set for population genomics. *Mol Ecol* 22:3124–3140.

116. Wu TD, Watanabe CK (2005) GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* 21:1859–1875.

117. Wu TD, Nacu S (2010) Fast and SNP-tolerant detection of complex variants and splicing in short reads. *Bioinformatics* 26:873–881.

118. Nei M (1987) *Molecular evolutionary genetics.* Columbia University Press, USA.

119. Holsinger KE, Weir BS (2009) Genetics in geographically structured populations: defining, estimating and interpreting F(ST). *Nat Rev Genet* 10:639–650.

120. Pritchard JK, Stephens M, Donnelly P (2000) Inference of population structure using multilocus genotype data. *Genetics* 155:945–959.

121. Falush D, Stephens M, Pritchard JK (2003) Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164:1567–1587.

122. Hubisz MJ, Falush D, Stephens M, Pritchard JK (2009) Inferring weak population structure with the assistance of sample group information. *Mol Ecol Res* 9:1322–1332.

123. Evanno G, Regnaut S, Goudet J (2005) Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Mol Ecol* 14:2611–2620.

124. Meirmans PG (2004) GENOTYPE and GENODIVE: Two programs for the analysis of genetic diversity of asexual organisms. *Mol Ecol Notes* 4: 792-794.

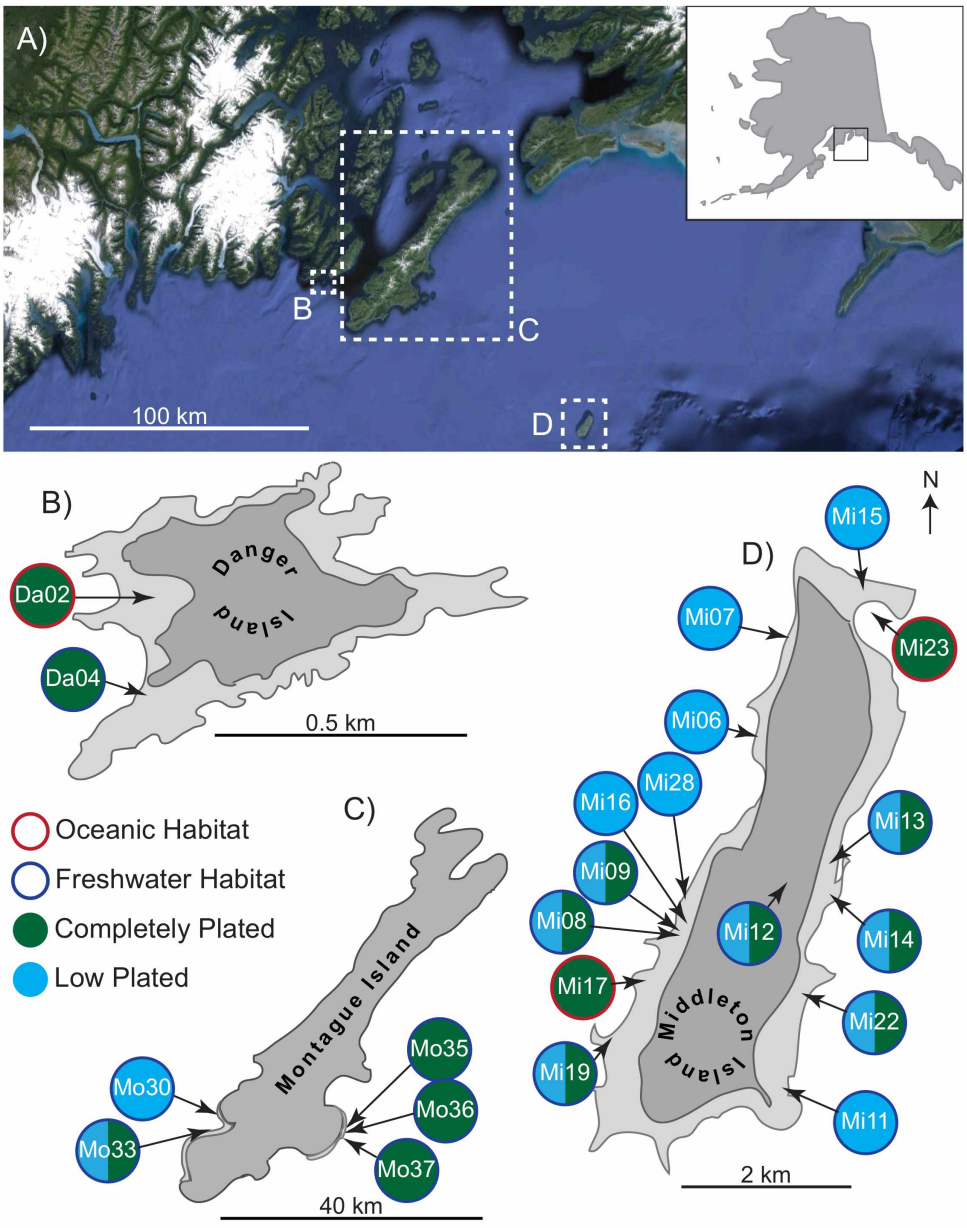

Figure 2.1. Map of sampling locations. A) Prince William Sound and the Gulf of Alaska, with Danger (B), Montague, (C), and Middleton (D) Islands outlined. Light gray areas represent the 1964 terraces on each island and dark gray areas indicate approximate pre-1964 sizes of each island. Inset: Alaska with box representing sampling location.

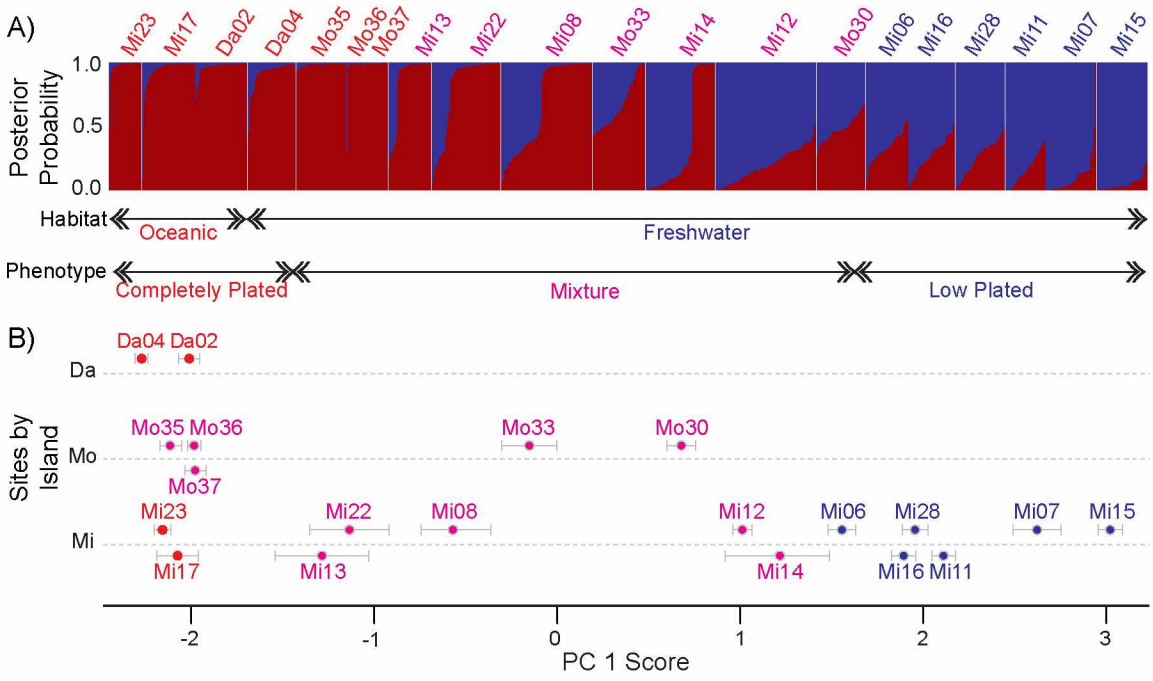

Figure 2.2. The major axis of genetic variation is a continuum of oceanic to freshwater genotypes. A) *STRUCTURE* analysis of the entire dataset reveals an optimum K of 2, representing oceanic (red) and freshwater (blue) genotypes. B) Distribution of mean scores for PC 1 +/- 1 SE. Principal component 1 accounts for 36% of the overall genetic variation.

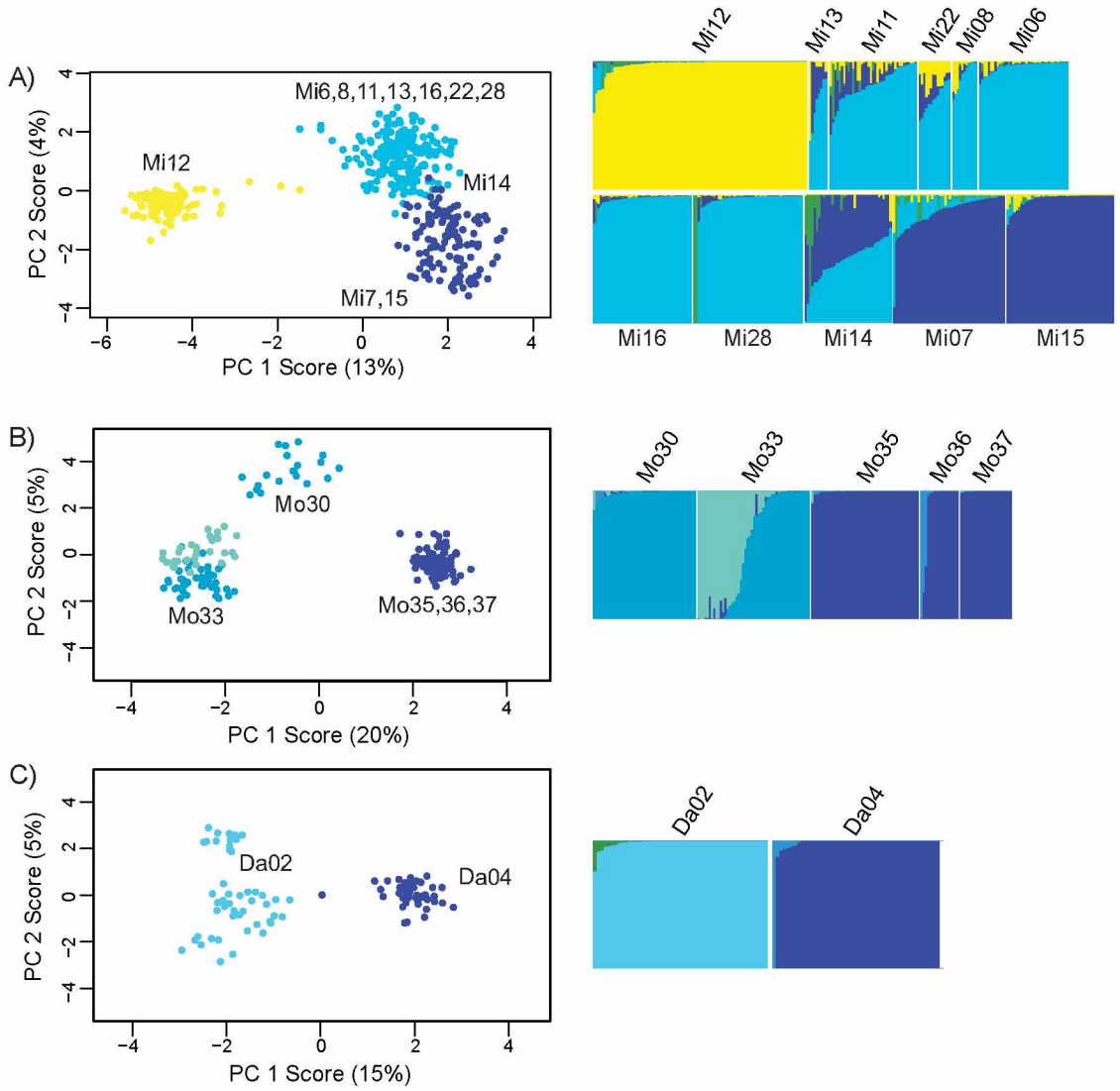

Figure 2.3. Plots of posterior probabilities of group assignment of each individual into clusters based upon the results of separate *STRUCTURE* analyses and corresponding principal components analyses color-coded by K means clustering assignment for each island. Percentages represent the amount of genetic variation explained by each PC. A) Post-1964 freshwater sites from Middleton Island separate into clusters representing sites 7, 14 and 15 and sites 6, 11, 13, 14, 16, 22, and 28. B) Montague sites separate into two major clusters that distinguish sites from the southeast (35, 36, and 37) and southwest (30 and 33). Site 33 contains two genotypic clusters. C) The freshwater (Da04) and oceanic (Da02) sites from Danger also form separate clusters.

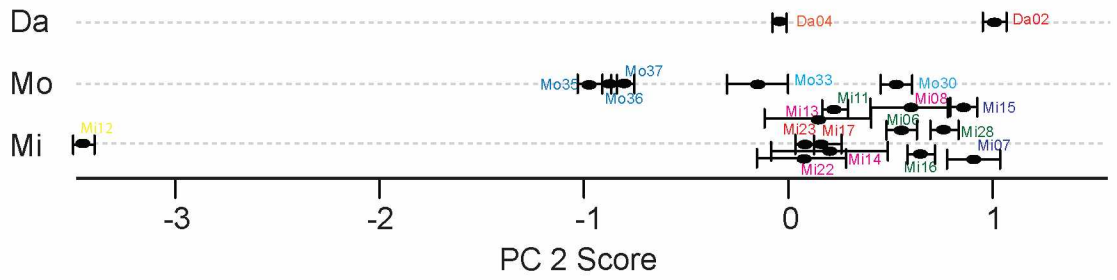

Figure 2.4. Mean PC 2 scores +/- 1SE. PC 2, which accounts for 16% of the overall genetic variation, separates the pre-1964 site from Middleton (Mi12) from the remaining sites.

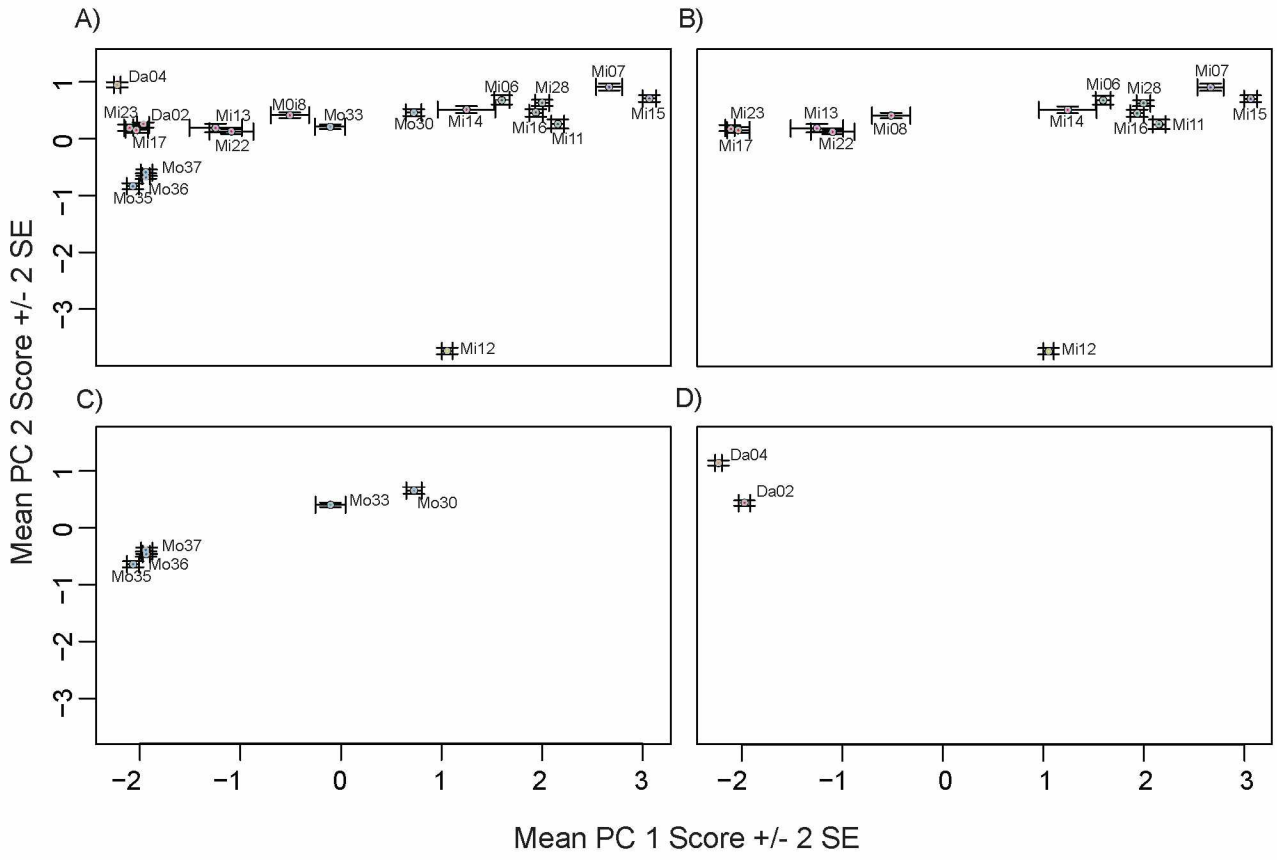

Figure 2.5. Mean PC 1 and 2 scores +/- 1SE for A) all sites analyzed, B) Middleton sites, C) Montague sites, and D) Danger sites.

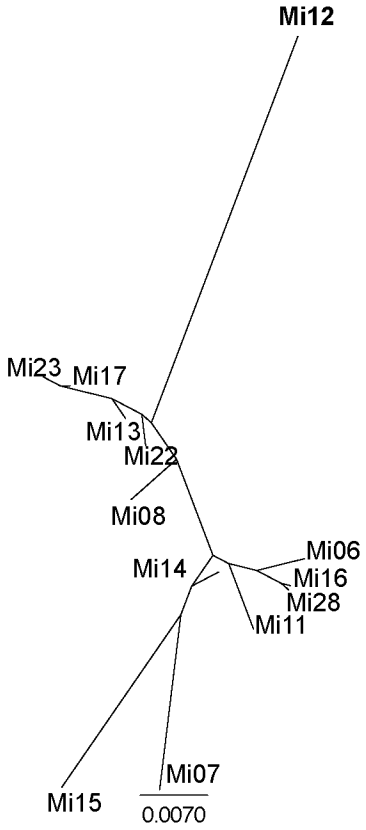

Figure 2.6. Neighbor joining tree of Middleton Island populations based on pairwise F_{ST}.

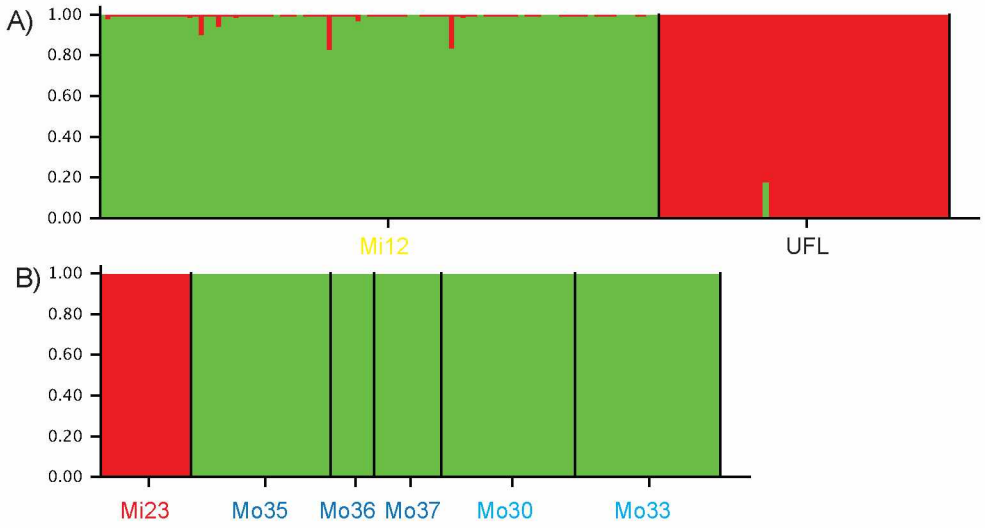

Figure 2.7. Plots of posterior probabilities of group assignment of each individual into clusters based upon the results of a *STRUCTURE* analysis. A) The pre-1964 site on Middleton (12) appears genetically distinct from the potential founding population from Upper Fire Lake (UFL). B) Adding an oceanic population to the Montague dataset shows clear separation between oceanic and Montague individuals, suggesting that there are two freshwater genotypes in Montague site 33.

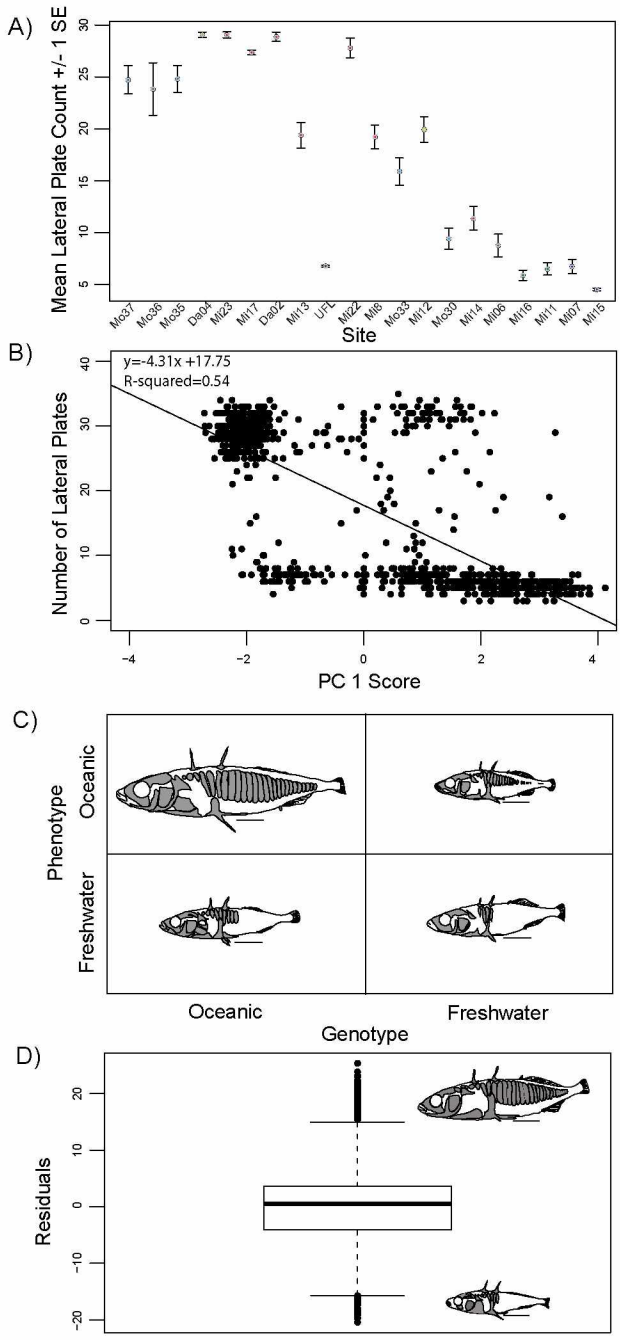

Figure 2.8. A) Mean lateral plate count for each site +/- 1 SE. B) There is a significant relationship between PC 1 score and lateral plate number. Each point represents an individual. C) Illustrations of representatives of each quadrant in panel B. Upper left is an individual with oceanic phenotype and genotype. Upper right is an individual with a freshwater genotype and oceanic phenotype. Lower right is an individual with a freshwater genotype and phenotype. Lower left is an individual with an oceanic genotype and freshwater phenotype. D) Boxplot of residuals with illustrations of individuals representing the strongest outliers.

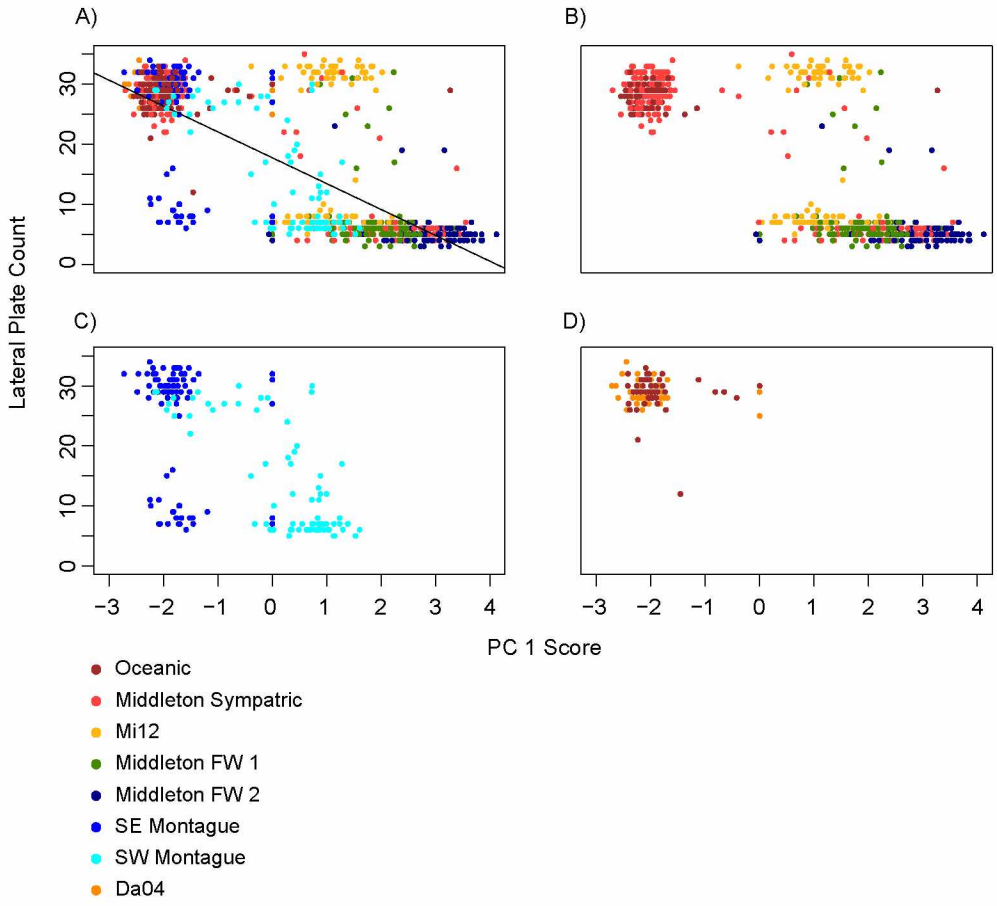

Figure 2.9. PC 1 score versus lateral plate number for individuals from A) all sites B) Middleton sites, C) Montague sites, and D) Danger sites. Each point is an individual color-coded by population.

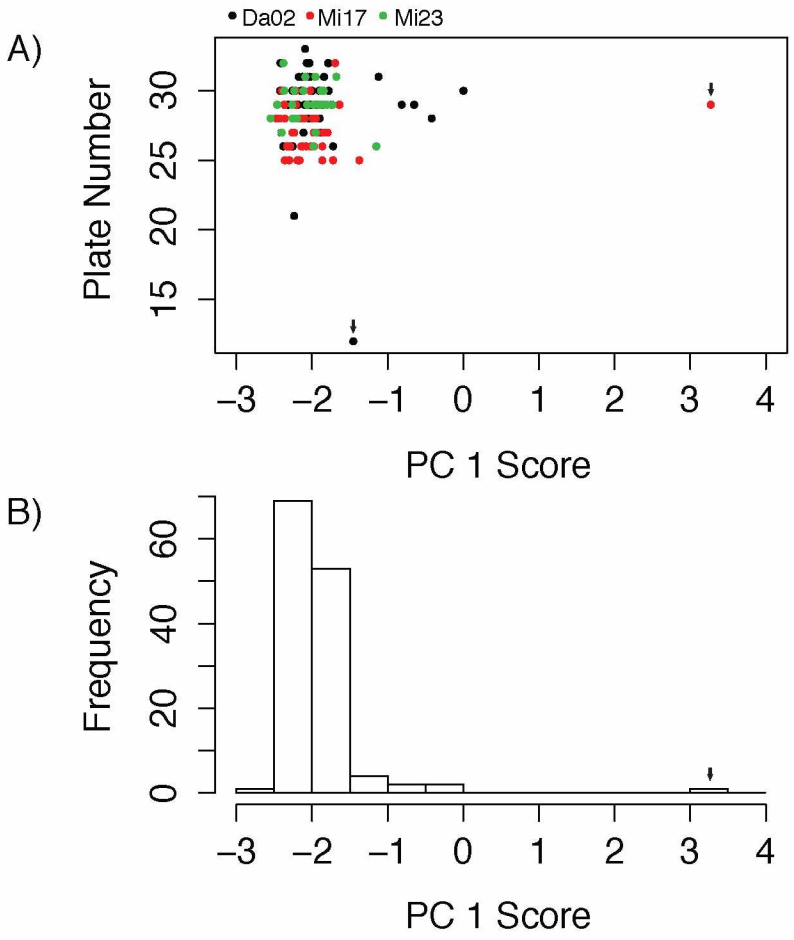

Figure 2.10. Most individuals caught in marine environments have oceanic phenotypes and genotypes, but freshwater phenotypes and genotypes appear to be maintained at low frequency. A) PC 1 score versus lateral plate number for oceanic sites Da02, Mi17, and Mi23. One individual from Mi17 has a freshwater genotype and one individual from Da02 has a freshwater phenotype. B) Frequency distribution of PC 1 scores in oceanic sites. One individual from Mi17 has a freshwater genotype.

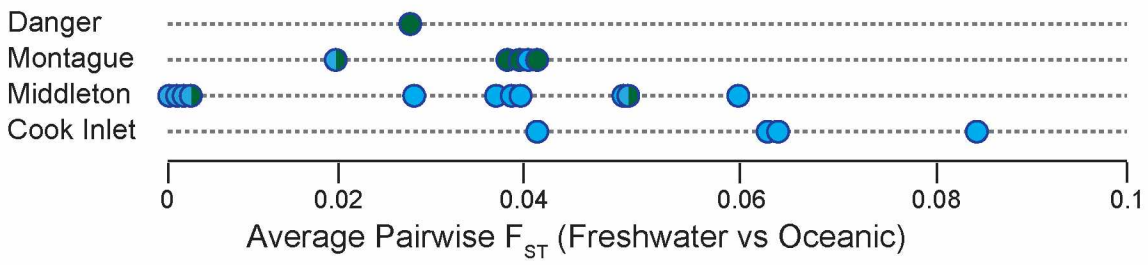

Figure 2.11. Comparison of pairwise F_{ST} between freshwater and oceanic populations from Cook Inlet and each island demonstrates that within just a few decades, freshwater populations from uplifted islands have genetically diverged from oceanic ancestors to nearly the same extent as seen in populations that were founded thousands of years ago. Data from Cook Inlet from Hohenlohe et al. (2010). Colors represent dominant lateral plate phenotype (green=high plated and blue=low plated).

Table 2.1. Sample collection data, including site name (Da=Danger Island, Mi=Middleton Island, Mo=Montague Island, UFL=Upper Fire Lake), coordinates, and habitat type based on water chemistry (FW = freshwater, OC = oceanic).

Site	Island/Region	Latitude	Longitude	Habitat
Da02	Danger	59.925	-148.083	OC
Da04	Danger	59.923	-148.078	FW
Mi06	Middleton	59.439	-146.331	FW
Mi07	Middleton	59.448	-146.325	FW
Mi08	Middleton	59.426	-146.357	FW
Mi09	Middleton	59.4282	-146.3532	FW
Mi11	Middleton	59.412	-146.338	FW
Mi12	Middleton	59.439	-146.331	FW
Mi13	Middleton	59.432	-146.314	FW
Mi14	Middleton	59.437	-146.311	FW
Mi15	Middleton	59.463	-146.299	FW
Mi16	Middleton	59.429	-146.349	FW
Mi17	Middleton	59.427	-146.359	OC
Mi19	Middleton	59.4183	-146.3696	FW
Mi22	Middleton	59.412	-146.333	FW
Mi23	Middleton	59.461	-146.296	OC
Mi28	Middleton	59.826	-147.478	FW
Mo30	Montague	59.823	-147.905	FW
Mo33	Montague	59.83705	-147.911	FW
Mo35	Montague	59.872	-147.438	FW
Mo36	Montague	59.872	-147.437	FW
Mo37	Montague	59.872	-147.437	FW
UFL	Cook Inlet	61.349	-149.536	FW

Table 2.2. Per sample read statistics after cleaning and demultiplexing.

Illumina Run	Lane	Length	Total Sequences	Low Quality	Ambiguous Barcodes	Ambiguous RAD Site	Retained Reads	Percent Retained
1014.120208_SN747_0217_BD0N93ACXX	6	101bp	182,762,273	20,088,037	18,799,510	18,237,584	125,637,142	68.74%
1244.120913_SN747_0261_BD1B3MACXX	4	101bp	131,141,192	9,613,822	6,269,459	6,181,216	109,076,695	83.18%
1245.120913_SN747_0261_BD1B3MACXX	6	101bp	206,462,539	32,097,222	8,155,160	7,795,305	158,414,852	76.73%
1246.120913_SN747_0261_BD1B3MACXX	7	101bp	154,194,316	22,701,413	13,257,570	7,513,548	110,721,785	71.81%
1355.130308_SN747_0284_BD1TWNACXX	6	101bp	204,867,187	14,081,086	97,750,540	5,010,495	88,025,066	42.97%
816.111021_SN747_0188_AC09HWACXX	1	101bp	134,207,441	8,531,479	9,659,014	10,662,855	105,354,093	78.50%
817.111021_SN747_0188_AC09HWACXX	2	101bp	100,624,872	4,801,129	8,530,767	8,690,723	78,602,253	78.11%
818.111021_SN747_0188_AC09HWACXX	3	101bp	167,850,213	14,141,932	18,824,350	10,860,644	124,023,287	73.89%
853.111021_SN747_0188_AC09HWACXX	6	101bp	128,494,316	8,192,987	10,455,374	7,035,635	102,810,320	80.01%
854.111104_SN747_0191_BD0GGAACXX	8	101bp	169,797,064	15,811,286	9,965,169	9,353,829	134,666,780	79.31%
855.111104_SN747_0191_BD0GGAACXX	7	101bp	156,398,983	12,059,427	12,264,617	10,517,437	121,557,502	77.72%
856.111122_SN747_0195_BD0F9KACXX	8	101bp	147,245,303	26,478,540	9,987,762	9,882,678	100,896,323	68.52%
Total			1,884,045,699	188,598,360	223,919,292	111,741,949	1,359,786,098	
Average			157,003,808	15,716,530	18,659,941	9,311,829	113,315,508	72.17%
Standard Error			9,114,296	2,360,023	7,277,543	973,423	6,167,810	

Table 2.3. Heat map of pairwise comparisons of genetic distance (F_{ST}) among sites. Da=Danger Island. Mi=Middleton Island. Mo=Montague Island. UFL=Upper Fire Lake. Dark green represents the least amount of genetic divergence, yellow represents moderate levels, and red represents the most differentiated sites.

Danger		Middleton Freshwater							Middleton Sympatric						Montague					Cook Inlet
	Da04	Mi06	Mi07	Mi11	Mi15	Mi16	Mi28	Mi12	Mi08	Mi13	Mi14	Mi22	Mi17	Mi23	Mo30	Mo33	Mo35	Mo36	Mo37	UFL
Da02	0.041	0.041	0.062	0.044	0.068	0.045	0.045	0.062	0.008	0.006	0.023	0.006	0.003	0.004	0.043	0.026	0.046	0.037	0.040	0.066
Da04		0.094	0.121	0.094	0.150	0.094	0.092	0.138	0.032	0.040	0.052	0.032	0.032	0.041	0.104	0.074	0.138	0.135	0.140	0.165
Mi06			0.036	0.016	0.040	0.009	0.009	0.069	0.013	0.025	0.010	0.021	0.034	0.038	0.035	0.031	0.090	0.071	0.080	0.100
Mi07				0.037	0.040	0.031	0.029	0.090	0.029	0.043	0.019	0.038	0.051	0.057	0.057	0.054	0.116	0.097	0.107	0.125
Mi11					0.038	0.013	0.014	0.058	0.016	0.027	0.010	0.022	0.038	0.041	0.030	0.030	0.090	0.072	0.081	0.095
Mi15						0.034	0.033	0.111	0.030	0.049	0.021	0.041	0.056	0.064	0.064	0.060	0.144	0.123	0.135	0.153
Mi16							0.002	0.067	0.016	0.029	0.010	0.025	0.038	0.042	0.034	0.033	0.090	0.071	0.080	0.097
Mi28								0.067	0.016	0.029	0.010	0.025	0.038	0.041	0.033	0.032	0.088	0.069	0.078	0.095
Mi12									0.038	0.046	0.045	0.041	0.050	0.055	0.081	0.070	0.134	0.110	0.121	0.152
Mi08										0.003	0.008	0.003	0.006	0.006	0.019	0.011	0.032	0.021	0.025	0.042
Mi13											0.012	0.001	0.003	0.004	0.029	0.015	0.041	0.030	0.034	0.057
Mi14												0.012	0.019	0.018	0.020	0.017	0.051	0.034	0.040	0.058
Mi22													0.003	0.003	0.025	0.013	0.034	0.023	0.027	0.046
Mi17														0.001	0.036	0.021	0.036	0.028	0.031	0.054
Mi23															0.040	0.022	0.044	0.039	0.041	0.066
Mo30																0.032	0.100	0.079	0.089	0.112
Mo33																	0.072	0.054	0.061	0.088
Mo35																		0.002	0.011	0.170
Mo36																			0.011	0.165
Mo37																				0.173

Table 2.4. Posterior probabilities of assignment to the oceanic and freshwater clusters in the overall *STRUCTURE* analysis.

Population	Oceanic Cluster	Freshwater Cluster	N
Mo37	0.992	0.008	23
Mo36	0.946	0.054	15
Mo35	0.987	0.013	48
Da04	0.911	0.089	46
Mi23	0.965	0.035	31
Mi17	0.935	0.065	50
Da02	0.967	0.033	49
Mi13	0.835	0.165	41
Mi22	0.792	0.208	65
Mi08	0.675	0.325	87
Mo33	0.668	0.332	50
Mi12	0.194	0.806	96
Mo30	0.481	0.519	46
Mi14	0.381	0.619	66
Mi06	0.328	0.672	40
Mi16	0.269	0.731	44
Mi28	0.257	0.743	47
Mi11	0.157	0.843	38
Mi07	0.107	0.893	48
Mi15	0.048	0.952	48

Table 2.5. Posterior probabilities of assignment to clusters in the Middleton freshwater *STRUCTURE* analysis. Individuals from Mi08, Mi13, Mi14, and Mi22 are those with freshwater phenotypes and genotypes (low lateral plate count and positive PC 1 scores).

Site	Cluster 1	Cluster 2	Cluster 3	Cluster 4	N
Mi06	0.016	0.965	0.004	0.015	40
Mi07	0.863	0.096	0.027	0.015	50
Mi08	0.026	0.893	0.004	0.007	11
Mi11	0.104	0.816	0.035	0.045	39
Mi12	0.002	0.013	0.009	0.976	96
Mi13	0.222	0.673	0.000	0.104	8
Mi14	0.432	0.487	0.069	0.011	39
Mi15	0.968	0.011	0.003	0.017	48
Mi16	0.017	0.971	0.005	0.007	44
Mi22	0.132	0.712	0.000	0.156	14
Mi28	0.016	0.937	0.044	0.003	49

Table 2.6. Posterior probabilities of assignment to the Middleton pre-1964 site (Mi12) and Upper Fire Lake (UFL) clusters in the *STRUCTURE* analysis.

Population	Mi12 Cluster	UFL Cluster	N
Mi12	0.993	0.007	96
UFL	0.005	0.995	50

Table 2.7. Results of Analyses of Molecular Variance (AMOVA). For each analysis, most of the variation is at the level of the individual, which is consistent with recent colonization. Habitat Type = freshwater or oceanic, based on water chemistry. Lateral Plate Morph represents the mean phenotype of the population (low, partially, or high plated).

Comparison	Source of Variation	Nested In	% Variation	F-Statistic	P Value
All Populations	Within Individual	—	73.6	F_it	—
	Among Individual	Population	4.3	F_is	<0.001
	Among Population	Island	15.3	F_sc	<0.001
	Among Region	—	6.8	F_ct	<0.001
All Populations	Within Individual	—	72.7	F_it	—
	Among Individual	Population	4.2	F_is	<0.001
	Among Population	Habitat Type	17.7	F_sc	<0.001
	Among Habitat Type	—	5.3	F_ct	0.003
All Populations	Within Individual	—	77.6	F_it	—
	Among Individual	Population	4.9	F_is	0.001
	Among Population	Lateral Plate Morph	12.7	F_sc	0.001
	Among Lateral Plate Morph	—	4.7	F_ct	0.001
Middleton	Within Individual	—	71.2	F_it	—
	Among Individual	Population	4.8	F_is	<0.001

Table 2.7 continued

	Among Population	Habitat Type	10.1	F_sc	<0.001
	Among Habitat Type	—	13.9	F_ct	<0.001
Middleton FW	Within Individual	—	79.8	F_it	—
	Among Individual	Population	5.3	F_is	<0.001
	Among Population	Watershed	6.2	F_sc	<0.001
	Among Watershed	—	9.4	F_ct	0.016
Montague	Within Individual	—	75.4	F_it	—
	Among Individual	Population	2.5	F_is	<0.001
	Among Population	Region	4.7	F_sc	<0.001
	Among Region	—	17.4	F_ct	<0.001
Danger	Within Individual	—	83.4	F_it	—
	Among Individual	Population	2.9	F_is	<0.001
	Among Population	—	13.7	F_st	<0.001
Ocean	Within Individual	—	89.1	F_it	—
	Among Individual	Population	5.6	F_is	<0.001
	Among Population	—	5.2	F_st	<0.001

Table 2.8. Summary genetic statistics and (standard deviation) for groups of sites split into those calculated for 1) only nucleotide positions that are polymorphic in at least one site ("Variant Positions"), or 2) all nucleotide positions across all RAD sites regardless of whether they are polymorphic or fixed ("All Positions"). These statistics include the average number of individuals genotyped at each locus (N), the number of variable sites unique to each site (Private), the number of polymorphic (top) or total (bottom) nucleotide sites across the data set (Sites), percentage of polymorphic loci (% poly), the average frequency of the major allele (P), the average observed heterozygosity per locus (Hobs), the average nucleotide diversity (π), and the average Wright's inbreeding coefficient (FIS). "Ocean" = Danger site Da02 and Middleton Mi17 and Mi23. "Middleton FW 1" = Mi7 and Mi15. "Middleton FW 2" = sites Mi06, Mi11, Mi16, and Mi28. "Middleton sympatric" = Mi13, Mi14, and Mi22. "Montague SE" = Mo30 and Mo33. "Montague SW" = Mo35, Mo36 and Mo37.

Variant Positions							
Site Group	N	Private	P	Hobs	π	F_{IS}	Sites
Ocean	43.26 (10.31)	3810 (1524)	0.9612 (0.0003)	0.0558 (0.0011)	0.0584 (0.0003)	0.0162 (0.0100)	123203 (67.51)
Danger FW	46.06	1594	0.970679	0.0386	0.0409	0.0065	123491
Middleton FW 1	47.99 (5.994)	956 (2378)	0.9618 (0.0055)	0.0503 (0.0104)	0.0544 (0.0112)	0.0159 (0.0058)	123428 (69.30)
Middleton FW 2	41.81 (8.573)	1116 (2496)	0.9557 (0.0074)	0.0602 (0.0094)	0.0646 (0.0123)	0.0173 (0.0145)	123328 (27.72)
Middleton Sympatric	53.05 (14.70)	3939 (1604)	0.9557 (0.0067)	0.0584 (0.0107)	0.0660 (0.0113)	0.0303 (0.0063)	123481 (116.0)
Middleton Site 12	93.92	515	0.963278	0.0467	0.0498	0.0094	123182
Montague SW	47.59 (36.84)	935 (425.1)	0.9561 (0.0053)	0.0604 (0.0100)	0.0631 (0.0105)	0.0090 (0.0056)	123379 (26.16)
Montague SE	29.09 (16.41)	353 (210.3)	0.9695 (0.0006)	0.0403 (0.0015)	0.0422 (0.0003)	0.0051 (0.0031)	123510 (6.658)
Upper Fire Lake	49.69	908	0.972678	0.0380	0.0375	0.0008	123505
All Positions							
Site Group	N	P	Hobs	π	F_{IS}	Sites	% Poly
Ocean	43.40 (10.33)	0.9981 (0.0000)	0.0027 (0.0001)	0.0029 (0.0000)	0.0008 (0.0005)	2500030 (67.86)	0.0493 (0.0000)
Danger FW	46.32	0.9986	0.0019	0.0020	0.0003	2500318	0.0494 (0.0000)
Middleton FW 1	48.30 (0.8282)	0.9981 (0.0001)	0.0025 (0.0002)	0.0027 (0.0001)	0.0008 (0.0004)	2500234 (36.77)	0.0494 (0.0000)

Table 2.8 continued

Middleton FW 2	42.06 (3.940)	0.9978 (0.0000)	0.0030 (0.0001)	0.0032 (0.0000)	0.0009 (0.0004)	2500155 (32.59)	0.0493 (0.0000)
Middleton Sympatric	94.70 (17.10)	0.9981 (0.0001)	0.0023 (0.0001)	0.0025 (0.0001)	0.0005 (0.0004)	2500238 (119.5)	0.0494 (0.0000)
Middleton Site 12	53.33	0.9978	0.0029	0.0033	0.0015	2500010	0.0493 (0.0000)
Montague SW	47.79 (2.897)	0.9978 (0.0001)	0.0030 (0.0002)	0.0031 (0.0001)	0.0004 (0.0002)	2500215 (23.33)	0.0493 (0.0000)
Montague SE	29.19 (16.46)	0.9985 (0.0000)	0.0020 (0.0001)	0.0021 (0.0000)	0.0003 (0.0002)	2500229 (5.508)	0.0494 (0.0000)
Upper Fire Lake	49.83	0.9987	0.0019	0.0019	0.0000	2500137	0.0494 (0.0000)

Table 2.9. Posterior probabilities of assignment to clusters in the Montague *STRUCTURE* analysis.

	Cluster			
Site	1	2	3	4
Mo30	0.003	0.008	0.007	0.982
Mo33	0.018	0.001	0.424	0.557
Mo35	0.994	0.003	0.003	0.001
Mo36	0.854	0.144	0.001	0.001
Mo37	0.997	0.001	0.002	0.001

Table 2.10. Posterior probabilities of assignment to the oceanic (Da02) and freshwater (Da04) clusters in the *STRUCTURE* analysis.

Population	Oceanic Cluster	Freshwater Cluster	N
Da02	1.0	0.0	49
Da04	0.033	0.967	47

Table 2.11. Summary genetic statistics for all sites calculated for 1) only nucleotide positions that are polymorphic in at least one site, or 2) all nucleotide positions across all RAD sites regardless of whether they are polymorphic or fixed. These statistics include the average number of individuals genotyped at each locus (N), percentage of polymorphic loci (% poly), the average frequency of the major allele (P), the average observed homozygosity per locus (Hobs), the average nucleotide diversity (π), and the average Wright's inbreeding coefficient (F_{IS}).

Site	N	% Poly	P	H_{OBS}	π	F_{IS}
Polymorphic in entire set of sites						
Da02	48.88	4.93	0.998	0.0028	0.0029	0.0004
Da04	46.32	4.94	0.999	0.0019	0.0020	0.0003
Mi06	39.74	4.93	0.998	0.0030	0.0032	0.0006
Mi07	48.89	4.93	0.998	0.0024	0.0026	0.0011
Mi08	36.66	4.93	0.998	0.0030	0.0032	0.0009
Mi11	38.01	4.93	0.998	0.0028	0.0032	0.0014
Mi12	94.70	4.94	0.998	0.0023	0.0025	0.0005
Mi13	40.88	4.93	0.998	0.0028	0.0031	0.0015
Mi14	71.03	4.93	0.998	0.0029	0.0034	0.0018
Mi15	47.72	4.94	0.998	0.0026	0.0027	0.0005
Mi16	43.74	4.93	0.998	0.0030	0.0032	0.0007
Mi17	49.83	4.93	0.998	0.0027	0.0029	0.0013
Mi22	64.77	4.92	0.998	0.0028	0.0032	0.0018
Mi23	31.48	4.93	0.998	0.0027	0.0029	0.0006
Mi28	46.75	4.93	0.998	0.0030	0.0032	0.0007
Mo30	45.74	4.94	0.998	0.0029	0.0030	0.0006
Mo33	49.84	4.93	0.998	0.0031	0.0032	0.0003
Mo35	47.86	4.94	0.998	0.0021	0.0021	0.0001
Mo36	16.77	4.94	0.999	0.0019	0.0021	0.0004
Mo37	22.95	4.94	0.998	0.0020	0.0021	0.0002
UFL	49.83	4.94	0.999	0.0019	0.0019	0.0000
Polymorphic in the focal site						
Da02		NA	0.960925	0.942946	0.00288227	0.000398734
Da04		NA	0.970679	0.961429	0.00202103	0.0003
Mi06		NA	0.955751	0.938661	0.00318626	0.0006
Mi07		NA	0.963029	0.952094	0.00262713	0.0011
Mi08		NA	0.954994	0.938724	0.00324849	0.0009
Mi11		NA	0.95576	0.942421	0.0032035	0.0014

Table. 2.11 Continued

Mi12		NA	0.963278	0.953304	0.00245927	0.0005
Mi13		NA	0.958213	0.943469	0.00313516	0.0015
Mi14		NA	0.952591	0.941718	0.00340623	0.0018
Mi15		NA	0.960585	0.94721	0.00274771	0.0005
Mi16		NA	0.95553	0.938751	0.0031903	0.0007
Mi17		NA	0.961095	0.9450	0.0028885	0.0013
Mi22		NA	0.95699	0.942448	0.00321066	0.0018
Mi23		NA	0.961447	0.944694	0.00286329	0.0006
Mi28		NA	0.955873	0.939268	0.00316933	0.0007
Mo30		NA	0.957201	0.941948	0.00303252	0.0006
Mo33		NA	0.955019	0.937346	0.00320108	0.0003
Mo35		NA	0.969021	0.95834	0.00209344	0.0001
Mo36		NA	0.970137	0.961306	0.00206855	0.0004
Mo37		NA	0.969212	0.95947	0.00208994	0.0002
UFL		NA	0.972678	0.962037	0.00185155	0.0000

Chapter 3 Ancient mtDNA Lineage Diversity in Threespine Stickleback Is Not Associated With Contemporary Patterns of Phenotypic or Nuclear Genetic Variation[1]

Abstract

Does deep evolutionary divergence influence contemporary patterns of variation? Two ancestral mitochondrial lineages of threespine stickleback (*Gasterosteus aculeatus*) co-occur in populations in the North Pacific Basin. These two clades are estimated to have diverged approximately 1 million years ago during a prolonged period of allopatry. Geographic separation could have led to reproductive barriers that persist between these lineages in contemporary populations, epistatic interactions among loci in the mitochondrial and nuclear genomes, or only neutral changes due to drift. We analyze these scenarios in stickleback populations from the Cook Inlet Basin and Middleton Island, Alaska that exhibit considerable spatial heterogeneity in clade proportions. We found no relationship between mtDNA lineage and a suite of armor traits or body shape. We also tested for nuclear genetic differentiation between the mtDNA lineages and examined genomically-localized patterns of variation. Our findings suggest that the two mtDNA clades are not spatially segregated or reproductively isolated, and that deep divergence does not have residual influence on contemporary nuclear genetic variation.

[1] Lescak, E.A., Wund, M., Bassham, S., Catchen, J., Colgren, J., Lucas, R.B., Prince, D.J., Sherbick, M.L., von Hippel, F.A., and Cresko, W.A. Prepared for submission to The Journal of Fish Biology.

Introduction

Determining how historical and contemporary forces shape variation in wild populations is a central focus of biogeography. Interpretation of spatial patterns in diversity should account for both microevolution, which produces population-level variation, and historical divergence of clades to determine how habitat-specific phenotypes evolve in lineages with different genetic backgrounds (Scribner et al. 2001). The advent of modern genomic techniques has allowed the examination of both deep and contemporary divergence in wild populations (e.g. Bernatchez et al. 2010; Hohenlohe et al. 2010; Deagle et al. 2012; Jones et al. 2012; Bourret et al. 2014; Brawand et al. 2014; Hellenthal et al. 2014; Vernot and Akey 2014). Studying individuals with divergent evolutionary histories can provide insights into how natural selection and ancestral divergence shape landscape-scale patterns of phenotypic and genetic variation (Ravinet et al. 2013).

Study of the threespine stickleback (*Gasterosteus aculeatus*; hereafter, 'stickleback') presents the opportunity to address this question. Oceanic stickleback have repeatedly and independently colonized newly available freshwater habitat throughout their holarctic distribution (Bell and Foster 1994), resulting in resident freshwater populations with widespread phenotypic variation that often evolves in parallel (Bell and Foster 1994; Hendry et al. 2013). Parallel evolution of body shape (Leinonen et al. 2006), armor phenotypes (Moodie and Reimchen 1976; Reimchen 1994), and trophic morphology (Hendry and Taylor 2004; Berner et al. 2008) have been documented in resident freshwater populations across the species distribution, including Alaska (Bell et al. 1993). Many of these phenotypes have a known genetic basis (Peichel et al. 2001; Colosimo et al.

2004, 2005; Cresko et al. 2004; Shapiro et al. 2004; Albert et al. 2008; Chan et al. 2010), suggesting that phenotypic divergence is due primarily to adaptation.

Individuals of this species complex descend from two ancient mitochondrial DNA (mtDNA) lineages that diverged during a period of allopatry approximately 1 MYA (Ortí et al. 1994; Nielsen and Wakeley 2001). The Euro North American clade (ENAC) is found in both the Atlantic and Pacific basins, while the Trans North Pacific Clade (TNPC) is restricted to the Pacific (Ortí et al. 1994; Johnson and Taylor 2004). The North Pacific Basin is a region of widespread clade admixture (Lescak et al. 2014). However, the causes of spatial heterogeneity in clade proportions in this region remain unknown (Johnson and Taylor 2004; Lescak et al. 2014).

A population-level association between reduced armor and presence of the TNPC has been documented in British Columbia (O'Reilly et al. 1993; Deagle et al. 1996) as well as the Seas of Okhotsk and Japan (Higuchi and Goto 1996). However, no association between morphology and clade was detected in analyses of additional populations from British Columbia or the Cook Inlet Basin, Alaska (Johnson and Taylor 2004). The lack of association could be a result of the wide variation in adaptive landscapes across these broad geographic ranges and strong relationship between environmental selection pressure and armor morphology (Johnson and Taylor 2004).

Our aim is to test whether ancient variation detected via mtDNA clade influences contemporary patterns of phenotypic or nuclear genetic variation. We hypothesized that the prolonged period of allopatry could have resulted in one or more scenarios. The two lineages may have become reproductively isolated, which would result in genome-wide nuclear genetic, and likely corresponding phenotypic, differences. Epistatic interactions

between the mitochondrial and nuclear genomes could have also evolved, which would be detectable through tests of genetic divergence between the two lineages at specific nuclear loci. Lastly, the period of allopatry could have been largely neutral, leaving a signature only in the non-recombining mtDNA. We tested for associations between mtDNA clade and phenotype in populations from the Cook Inlet Basin and Middleton Island, Alaska, which are both regions of clade admixture (Lescak et al. 2014), using both armor traits and whole body shape. In the Middleton Island populations, we also tested for significant partitioning of nuclear genetic variation by clade, and seek to identify genomic regions of significant divergence between the two lineages.

Methods

Field collections. Collections were made between 2005 and 2011 from locations shown in Figure 3.1 (see Table A.1 for site coordinates). Stickleback were collected using 0.32 and 0.64 cm mesh minnow traps set near shore and left overnight. Fish were euthanized with an overdose of neutral pH MS-222 anesthetic and preserved in 95% ethanol. We sampled six resident freshwater populations from the Cook Inlet Basin and 11 resident freshwater and two oceanic populations from Middleton Island (Fig. 3.1). Freshwater and oceanic sites were designated based on salinity (Lescak et al. unpublished data). Because the Barabara population, located on the southeast shore of Cook Inlet, is morphologically unusual (see results), two nearby populations (Two Island Lake and Sportfish Lake) for which we had already collected samples were also analyzed to determine whether traits found in the Barabara population are typical of lakes in the watershed. Fish from these lakes had previously been fixed in 10% buffered formalin and stained with Alizarin Red S, while fin clips had been collected for genetic analyses from

different individuals. Thus, because morphological and genetic samples of Two Island and Sportfish Lakes were not from the same individuals, these populations were not included in individual-level analyses relating clade identity to morphology.

Sample preparation. Caudal and pectoral fins were clipped for DNA extraction. Each set of fins and corresponding body was assigned a unique identification number for association of genotype with phenotype. After fins were clipped, bodies were fixed in 10% neutral buffered formalin for at least 48 hours, bleached in a 0.05% hydrogen peroxide solution, stained in a 0.1% Alizarin red S solution, destained in 1% KOH, and preserved in 70% ethanol. DNA was extracted from clipped fins using a Qiagen DNeasy kit and samples were normalized to 25 ng/μL (or 10 ng/ μL in populations with low yields).

Clade identification. The following primer pair was used to amplify a region of the cytochrome b gene: 14372 (5'-ATGGCAAGCCTACGAAAAACGCAC-3') and 15100 (5'-TGCTAGGGATGTAAGGGCAATTAG-3'; Lescak et al. 2014). Amplification reaction conditions were: 1x (2 min at 94°C), 28x (30 sec at 94°C, 15 sec at 58°C, 45 sec at 72°C), 1x (7 min at 72°C) and 1x (6 min at 94°C), 37x (60 sec at 94°C, 60 sec at 63°C, 120 sec at 72°C), and 1x (5 min at 72°C).

To assign the amplified cytochrome b gene fragments to either the TNPC or ENAC, we performed restriction digests with *BstXI* and *NlaIII* and size separated the restriction digest products on agarose gels. The amplified cytochrome b fragment in the TNP clade lacks *BstXI* recognition sites and includes one *NlaIII* site, while descendants of the ENA lineage have one *BstXI* recognition site and two *NlaIII* sites (Ortí et al., 1994). We performed independent digests with each of the two enzymes to increase confidence in clade assignment. All individuals were unambiguously assigned to one or the other lineage based

on the diagnostic restriction sites. We also sequenced the amplified gene fragment in a subset of 11 individuals (6 ENA and 5 TNP) from six Alaskan populations to confirm consistent correspondence between restriction site arrangement and mtDNA clade affiliation (GenBank accession numbers KM508783-KM508793).

Armor analysis. Pelvic spines were measured from the anterior-most point of the articulation with the pelvic girdle to the tip of the spine. Pelvic scores were assigned according to methods described in Bell et al. (1993). Lateral plates were counted visually from photographs. All other linear measurements were made from photographs using ImageJ v1.43r. Dorsal spines were measured from the anterior side of the articulation to the tip of the spine. The most anterior supporting plate that articulated with the second pterygiophore was measured for maximum height and total area. However, in 12 samples, no plates met this criterion. In these cases, the plate that came closest to articulation was measured. Morphometric data were size-standardized following Berner et al. (2011; Fig. 3.2).

Pearson correlations were performed between trait means and proportion TNPC for each population. These analyses were performed with and without Barabara, since this population was an outlier (Figs. 3.1, 3.3; Tables 3.1-3.2). A principal components analysis (PCA) was performed using the pcaMethods package (Stacklies et al., 2007) in R (R Core Team, 2014) to determine the overall distribution of phenotypic variation. Separate analyses of variance (ANOVAs) were performed on the first three principal components to test for significant differences in phenotype between clades, sexes, and regions (Mainland versus Middleton Island), and among populations. Tukey's post-hoc tests were used to determine which populations significantly differ.

Body shape analysis. Body shape was characterized using geometric morphometrics (Zelditch et al., 2004). Nineteen homologous landmarks were identified on 712 individuals (Fig. 3.2) and placed on photographs of the left lateral view using TPSdig v. 2.10 (Rohlf, 2010). Prior to geometric morphometric analysis of shape, the "unbend specimens" algorithm in TPSutil v. 1.47 (Rohlf, 2010) was used to remove non-biological shape differences due to specimen bending. Landmark coordinates were adjusted such that landmarks 1, 8 and 19 (Fig. 3.2) fell along a straight line, as they would naturally.

Using MorphoJ v. 1.05f (Klingenberg, 2011), generalized Procrustes analysis superimposition was performed on all specimens to obtain adjusted landmark coordinates that were independent of size, rotation and translation (Zelditch et al., 2004). To examine shape differences between regions (Cook Inlet versus Middleton Island), among populations, and between sexes and clades, canonical variates analysis (CVA) was performed in MorphoJ, which also produced deformations associated with principle axes of discrimination. Prior to CVA, shape differences due to body size (measured as centroid size, the sum of the squared distances between each landmark and the landmark configuration centroid) were removed using pooled within-group regression to account for possible slope heterogeneity among populations (Reist, 1986). Residuals from this regression were then used in CVA.

ANOVAs on CV scores were performed to determine whether shape differed between regions, ecotypes, sexes, and clades and among populations nested within regions. All factors except population were included as fixed factors in ANOVA models. Population was included as a nested (Mainland versus Middleton Island), random factor. Two separate CVAs (and subsequent ANOVAs) were performed. The first included all populations, and

determined whether shape differed between regions, ecotypes and sexes and among populations. The second CVA also included clade as a factor, and omitted the two populations for which clade information is not associated with individual specimens (Two Island and Sportfish Lakes).

Genetic sex determination. Stickleback that could not visually be assigned to a sex were genotyped using a PCR assay (Griffiths et al., 2000). Males show a sex-specific band at 371bp when amplified with the Ga1 primer pair. The PTC-200 thermocycler denatured the reactions at 94°C for 3 min and then performed 36 cycles of 94°C for 45s, 44°C for 45s, and 72°C for 45s before a final run of 72°C for 10 min.

RAD library preparation and sequence analysis. Genomic DNA from each Middleton Island individual was digested with the restriction enzyme SbfI-HF (NEB), and RAD-seq libraries were created as in Hohenlohe et al. (2010). Uniquely barcoded samples representing 76 to 96 individuals were run per lane in 12 total lanes of sequencing on an Illumina HiSeq 2000 platform. Read lengths were 101 nucleotides long including 6 nucleotide in-line barcodes. Raw sequence data were demultiplexed and filtered for quality using the process_radtags program in the *Stacks* software suite (Catchen et al. 2011, 2013). Reads were aligned against the stickleback reference genome (version BROADs1, Ensembl release 64) using GSnap (Wu and Watanabe 2005), allowing for up to 5 mismatches and gaps of length 2, disabling terminal alignments, and requiring unique alignments. The alignments were processed and genotypes were called for each locus across all samples using the pstacks, cstacks, and sstacks programs from *Stacks*. For a locus to be included in further analysis, we required that it be present in all compared populations and successfully genotyped in at least 75% of individuals from each population.

Genetic analysis. Population genetic statistics (major allele frequency, percent polymorphic loci, nucleotide diversity (π), and Wright's F statistics F_{IS} and F_{ST}) were calculated for every SNP using the populations program in *Stacks* (Catchen et al. 2011, 2013). To analyze population structure using multilocus genotypes, we used the populations program in *Stacks* to output filtered SNP data from all RAD loci across all populations into a file formatted for *STRUCTURE* (Pritchard et al. 2000; Falush et al. 2003; Hubisz et al. 2009). Because of computational limitations, we randomly chose three subsets of 1,000 SNPs each to complete the analysis. These three subsets were found to yield comparable results, so results from only one subset are included. For all analyses, 10,000 burn-in steps and 10,000 replicates were used, with 10 replicates for each potential *K* (number of genotypic groups). The optimal *K* for each analysis was chosen using the deltaK method (Evanno et al. 2005). This same set of SNPs was used in *GenoDive* (Meirmans and van Tiendren 2004) to identify the major axes of genetic variation using principal components analyses (PCA) and to test for significant partitioning of variation using an analysis of molecular variance (AMOVA).

Results

Alaskan stickleback populations were genetically diverse. The Cook Inlet Basin and Middleton Island were regions of mtDNA clade admixture. While the ENAC predominated over the TNPC overall (77% to 23%), TNPC proportions range from 0% to 100% among populations (Fig. 3.1). Populations from Middleton Island exhibited a great deal of nuclear genetic variation (Lescak et al. unpublished data) and significantly differed in scores from PC 1 ($F=115.2$, $p<0.001$), PC 2 ($F=655.7$, $p<0.001$), and PC 3 ($F=75.4$, $p<0.001$) in the analysis of 1000 loci.

Phenotype was not significantly associated with mtDNA clade. At the population level, correlations between mean armor traits and TNPC proportion were not significant, with the exception of right and left pelvic spine length in the analysis of the full dataset (Fig. 3.3; Table 3.2-3.3). However, when Barabara was removed from the analysis, due to it being the only site apparently fixed for the TNPC as well as the only site with pelvic reduction, the correlations were no longer significant. ANOVAs of scores from PC's 1-3 in the analysis of armor traits did not reveal any significant differences between sexes or clades in individuals from Cook Inlet, but did detect significant differences among sites ($p<0.001$; Figs. 3.4-3.5). Individuals from Middleton Island significantly differed in scores from PC's 1-3 based on sex and site ($p<0.001$; Fig. 3.5). PC 3 scores also significantly differed between clades, but this is likely due to differences in clade frequencies among sites (Fig. 3.1). This PC accounts for <1% of the overall phenotypic variation. Lateral plates had the highest loading on PC 1, plate height had the highest loading on PC 2, and pelvic spine lengths had the highest loadings on PC 3 (Table 3.1).

In analysis of body shape, significant differences in CV 1 were found between regions and sexes and among populations ($p<0.001$; Fig. 3.6; Table 3.4). Significant differences in CV 2 were found among populations and between ecotypes and sexes ($p<0.001$; Fig. 3.6; Table 3.4). In the CVA in which Two Island and Sportfish Lakes were removed due to lack of corresponding genetic and phenotypic data, significant differences in CV 1 were detected between regions and sexes and among populations ($p<0.001$; Fig. 3.6; Table 3.4). Significant differences in CV 2 scores were detected among populations and between ecotypes, sexes, and clades ($p<0.001$; Fig. 3.6; Table 3.4).

Nuclear genetic variation was not partitioned by mtDNA clade. The two lineages significantly differed in scores from PC 1 ($F=11.0$, $p<0.001$), but not PC 2 ($F=0.51$, $p=0.599$) or PC 3 ($F=0.27$, $p=0.766$) in the analysis of 1000 nuclear loci. PC 1, however, also had a significant population by clade interaction ($F=3.49$, $p<0.001$), suggesting that population-level variation drives differences between clades (Fig. 3.7). In support of this, global pairwise F_{ST} between the two lineages was 0.011 and nuclear genomic variation was not associated with mtDNA lineage (Tables 3.5-3.6). Analysis of molecular variance revealed that most of the variation is at the level of the individual (91%) and no variation is attributed to differences between clades. Analysis of Mi7, Mi14, and Mi16, which contained comparable proportions of each clade, reveal no intra-population structure by mtDNA lineage (Fig. 2.2-2.3) and had low levels of genetic divergence between individuals belonging to each clade (pairwise F_{ST} between the two lineages ranges from 0.001-0.080; Table 3.7). Scans of F_{ST} across the genome did not reveal any genomic regions that are divergent between the two lineages across all populations pooled or within populations (Fig. 3.8).

Discussion

Southcentral Alaska is a region of clade admixture. Our fine-scale sampling revealed a great deal of spatial heterogeneity in clade proportions over a relatively small geographic area. While lineage sorting would ultimately tend to result in one haplotype becoming fixed, this process takes time. It is possible that given the relatively young ages of our freshwater populations (approximately 13,000 years for the Cook Inlet Basin populations and <50 years for those on Middleton Island; Bell and Foster, 1994; Gelmond et al. 2009), not enough time has passed to allow for complete lineage sorting.

Barabara Lake is the first documented population in southcentral Alaska that is fixed for the TNPC (O'Reilly et al., 1993; Ortí et al., 1994; Deagle et al., 1996; Cresko, 2000; Hendry et al., 2002; Johnson and Taylor, 2004; Drevecky, 2011; Weigner, 2012). All populations previously surveyed in the Cook Inlet Basin had proportions of the TNPC of less than 25% (Ortí et al., 1994; Cresko, 2000) and only one out of 46 sampled populations in the Queen Charlotte Islands, British Columbia (Rouge Lake) was fixed for the TNPC (O'Reilly et al., 1993; Deagle et al., 1996; Johnson and Taylor, 2004). It should be noted that while Ortí et al. (1994) reported two populations in the Matanuska-Susitna Valley, Alaska that only contained the TNPC, the sample size was only one fish for each population; one of these populations was later shown not to be fixed (75% TNPC; Drevecky, 2011), while the other has not been resampled.

Typically, about 5% of anadromous stickleback in southcentral Alaska are of the TNPC (Cresko, 2000; Drevecky, 2011; Weigner 2012; Lescak et al. 2014). Since oceanic fish found freshwater populations, the expectation is that freshwater clade proportions would be comparable to those seen in oceanic populations. However, the Barabara, Horsehead and Grant populations show substantially higher proportions of the TNPC. Nearly half of the stickleback in Horsehead Lake were of the TNPC, despite the fact that the lake appears to open directly to the ocean every 6-9 years (MS Christy, personal communication) and would therefore be expected to be regularly colonized primarily by fish of the ENAC. Weigner (2012) found similarly high frequencies of TNPC in resident freshwater populations from the Bering Glacier region of Alaska.

It is unclear what could cause a population to become fixed for the rarer of the two lineages. It has been proposed that populations with a TNPC majority could be relicts that

colonized glacial refugia during one of several deglaciation events during the Pleistocene (O'Reilly et al., 1993), or that clade proportions are significantly associated with elevation (Johnson and Taylor, 2004). However, these hypotheses have been countered due to the lack of association between clade proportion and geographic location (Ortí et al., 1994; Deagle et al., 1996; Cresko, 2000). Alternatively, Drevecky (2011) suggested that selection acting on any of the mitochondrial genes might explain the persistence of the TNPC, since all genes located on the mitochondrial chromosome are linked due to a lack of recombination.

Phenotype was independent of mtDNA clade. Previous studies have shown that populations with higher proportions of the TNPC have greater reduction of body armor than populations with higher proportions of the ENAC (O'Reilly et al., 1993; Deagle et al., 1996). The Barabara population, which was fixed for the TNPC, showed the greatest reduction in body armor, including pelvic reduction (Table 3.2). However, similar to findings made by Johnson and Taylor (2004), morphological differences were not found between the clades in the admixed populations, suggesting that reduction in armor is not consistently associated with clade identity. The correlation that had previously been detected between clade and phenotype was at the level of the population and therefore not a causal connection of mtDNA variation directly affecting trait variation.

The lack of association between mtDNA clade and phenotypic variation confirms what is already well known about the stickleback radiation: morphological diversity is extensive among freshwater populations and relates to local adaptation (reviewed in Bell and Foster, 1994). This is best demonstrated by the distinct morphology of the Barabara Lake population, in which individuals are almost as large as the giant stickleback of British

Columbia (Reimchen, 1988) and exhibit pelvic reduction, in contrast to its close neighbors, which have a body size that is more typical of resident freshwater populations and full pelvic armor. While ecological data were not collected for each of our sample sites, it is well documented that predation regime and water chemistry influence the phenotypes of stickleback populations (reviewed in Bell and Foster, 1994). It is likely that some combination of these selective agents shape morphological variation in the populations surveyed. Ecological data were collected for populations from Middleton Island; significant variation was not observed among freshwater sites in a suite of water chemistry variables and they all have the same predation regimes (benthic macroinvertebrates and avian predators). Morphological variation among these sites is likely due to extent of secondary contact with oceanic ancestors (Gelmond et al., 2009).

Nuclear genetic variation was not partitioned by mtDNA lineage. Individuals belonging to the two ancient lineages appear to freely interbreed, based on low F_{ST} and lack of grouping by clade in *STRUCTURE* analyses. The significant difference we detected in PC 1 genetic scores between lineages is driven by a clade by population interaction. The absence of localized regions of divergence across the genome in comparisons of the two lineages suggests a lack of cytonuclear disequilibrium, which is also supported by the absence of a relationship between mtDNA lineage and microsatellite variation (Cresko, 2000). If the nuclear genome diverged concurrently with the mitochondrial genome approximately 1 MYA, the signal of divergence appears to have been lost, likely due to recombination in the nuclear genome and a prolonged period of secondary contact between the two clades.

Discordant patterns of genetic divergence between the mtDNA and nuclear DNA have been reported in other species. For example, a prolonged period of allopatry

produced an mtDNA cline in the barnacle *Notochthamalus scabrosus* off the coast of Chile that was not detectable in nuclear DNA (Zakas et al., 2014). Similarly, mtDNA, but not microsatellite, variation is significantly partitioned by geography in populations of Pacific herring (*Clupea palasii*) that were historically isolated during the Pleistocene (Liu et al., 2012). In the rock pocket mouse (*Chaetodipus intermedius*), concordance was not detected between mtDNA and pelage coloration, suggesting that this phenotype has evolved variation relatively recently and in response to local habitat conditions (Hoekstra et al., 2005). In contrast, concordant patterns of mitochondrial, nuclear, and phenotypic variation were observed in sister butterfly taxa *Anartia fatima* and *Anartia amathea* within their hybrid zone, supporting assortative mating within taxa and epistasis between mitochondrial and nuclear loci (Dasmahapatra et al. 2002).

Regions of clade admixture provide an opportunity to study introgression of distinct genomes following secondary contact. Our research highlights the importance of fine-scale sampling when examining genetic trends over large geographic regions. A lack of significant association between clade and morphology or nuclear genetic variation indicates that local adaptation is not constrained by the historical contingency of clade identity. Additional research is necessary to determine the reasons behind the coexistence of the clades, but given the lack of relationship between clade identity and morphological divergence or nuclear genetic structure, neutral processes are likely responsible for the observed spatial heterogeneity in mtDNA. The processes that created these two lineages did not lead to the evolution of reproductive isolation or epistatic selection in stickleback. Variation in mtDNA in this species represents ancient variation that has likely been erased in the nuclear genome due to recombination and widespread secondary contact.

Acknowledgements

We thank B.R. Harrison for productive discussions, H.L. Weigner for assistance in developing our genetic assays, and M.S. Christy for logistical support. This research was supported primarily by National Science Foundation DEB 0949053 to WAC and DEB 0919234 to FAvH, as well as IOS 102728 (WAC). Additional support came from a University of Alaska Anchorage Faculty Development Grant (FAvH), NIH grant 1R24GM079486-01A1 (WAC), NIH NRSA Ruth L. Kirschstein fellowship F32GM095213-01 (JMC), the M. J. Murdock Charitable Trust (WAC), and funds from The College of New Jersey (MAW). EAL was supported with funds from the University of Alaska Center for Global Change and Arctic Systems Research, the University of Alaska Anchorage, LGL Limited, the Rasmuson Fisheries Board, the American Fisheries Society, and the NSF Alaska EPSCoR Landscape Genetics program. All procedures were approved by the IACUC from Clark University (protocol #0010R), The College of New Jersey (protocol #1002-001), and the University of Alaska Anchorage (protocol #159870, 2009wund1, and 160406-1). Stickleback were collected in accordance with Alaska Department of Fish and Game permit numbers SF2009-016, SF2009-038, SF2009-065, SF2010-029, SF2010-111, SF2011-067 and SF2011-153.

References

Albert, A. Y., Sawaya, S., Vines, T. H., Knecht, A. K., Miller, C. T., Summers, B. R., Balabhadra, S., Kingsley, D. M. & Schluter, D. (2008). The genetics of adaptive shape shift in stickleback: pleiotropy and effect size. *Evolution* **62**, 76-85.

Bell, M. A. & Foster, S. A. (1994). The Biology of the Stickleback. Oxford University Press.

Bell, M. A., Ortí, G., Walker, J. A., & Koenings, J. P. (1993). Evolution of pelvic reduction in threespine stickleback fish: a test of competing hypotheses. *Evolution* **47**, 906-914.

Bernatchez, L., Renaut, S., Whiteley, A. R., Derome, N., Jeukens, J., Landry, L., Lu, G., Nolte, A. W., Ostbye, K., Rogers, S. M. & St-Cyr, J. (2010). On the origin of species: insights from the ecological genomics of lake whitefish. *Philosophical Transactions of the Royal Society of London B Biological Sciences* **365**, 1783-1800.

Berner, D., Adams, D. C., Grandchamp, A. –C., & Hendry, A. P. (2008). Natural selection drives patterns of lake-stream divergence in stickleback foraging morphology. *Journal of Evolutionary Biology* **21,** 1653-1665.

Berner, D., Kaeuffer, R., Grandchamp, A. –C., Raeymaekers, J. A. M., Rasanen, K., & Hendry, A. P. (2011). Quantitative genetic inheritance of morphological divergence in a lake-stream stickleback ecotype pair: implications for reproductive isolation. *Journal of Evolutionary Biology* **24,** 1975-1983.

Bourret, V., Dionne, M. & Bernatchez, L. (2014). Detecting genotypic changes associated with selective mortality at sea in Atlantic salmon: polygenic multilocus analysis surpasses genome scan. *Molecular Ecology* **23**, 4444-4457.

Brawand, D., Wagner, C. E., Li, Y. I., Malinsky, M., Keller, I., Fan, S., Simakov, O., Ng, A. Y., Lim, Z. W., Bezault, E., Turner-Maier, J., Johnson, J., Alcazar, R., Noh, H. J., Russell, P., Aken,

B., Alfoldi, J., Amemiya, C., Azzouzi, N., Baroiller, J. F., Barloy-Hubler, F., Berlin, A., Bloomquist, R., Carleton, K. L., Conte, M. A., D'Cotta, H., Eshel, O., Gaffney, L., Galibert, F., Gante, H. F., Gnerre, S., Greuter, L., Guyon, R., Haddad, N. S., Haerty, W., Harris, R. M., Hofmann, H. A., Hourlier, T., Hulata, G., Jaffe, D. B., Lara, M., Lee, A. P., MacCallum, I., Mwaiko, S., Nikaido, M., Nishihara, H., Ozouf-Costaz, C., Penman, D. J., Przybylski, D., Rakotomanga, M., Renn, S. C., Ribeiro, F. J., Ron, M., Salzburger, W., Sanchez-Pulido, L., Santos, M. E., Searle, S., Sharpe, T., Swofford, R., Tan, F. J., Williams, L., Young, S., Yin, S., Okada, N., Kocher, T. D., Miska, E. A., Lander, E. S., Venkatesh, B., Fernald, R. D., Meyer, A., Ponting, C. P., Streelman, J. T., Lindblad-Toh, K., Seehausen, O. & Di Palma, F. (2014). The genomic substrate for adaptive radiation in African cichlid fish. *Nature* **513**, 375-381.

Catchen, J., Hohenlohe, P. A., Bassham, S., Amores, A. & Cresko, W. A. (2013). Stacks: an analysis tool set for population genomics. *Molecular Ecology* **22**, 3124-3140.

Catchen, J. M., Amores, A., Hohenlohe, P., Cresko, W. & Postlethwait, J. H. (2011). Stacks: building and genotyping Loci de novo from short-read sequences. *G3* **1**, 171-182.

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Shapiro, M. D., Brady, S. D., Southwick, A. M., Absher, D. M., Grimwood, J., Schmutz, J., Myers, R. M., Petrov, D., Jonsson, B., Schluter, D., Bell, M. A. & Kingsley, D. M. (2010). Adaptive Evolution of Pelvic Reduction in Sticklebacks by Recurrent Deletion of a Pitx1 Enhancer. *Science (New York, NY)* **327**, 302-305.

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J., Schmutz, J., Myers, R. M., Schluter, D. & Kingsley, D. M. (2005). Widespread parallel

evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science (New York, NY)* **307**, 1928-1933.

Colosimo, P. F., Peichel, C. L., Nereng, K., Blackman, B. K., Shapiro, M. D., Schluter, D. & Kingsley, D. M. (2004). The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology* **2**, E109.

Cresko W. A. (2000) *The ecology and geography of speciation: a case study using an adaptive radiation of threespine stickleback in Alaska*. PhD Thesis, Clark University, Worcester, MA.

Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P., Bell, M. A., Kimmel, C. B. & Postlethwait, J. H. (2004). Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences of the United States of America* **101**, 6050-6055.

Dasmahapatra, K. K., Blum, M. J, Aiello, A., Hackwell, S., Davies, N., Bermingham, E. P, & Mallet, J. (2002). Inferences from a rapidly moving hybrid zone. *Evolution* **56,** 741-753.

Deagle, B. E., Jones, F. C., Chan, Y. F., Absher, D. M., Kingsley, D. M. & Reimchen, T. E. (2012). Population genomics of parallel phenotypic evolution in stickleback across stream-lake ecological transitions. *Proc Biol Sci* **279**, 1277-1286.

Deagle B.E., Reimchen, T.E. & Levin, D.B. (1996) Origins of endemic stickleback from the Queen Charlotte Islands: mitochondrial and morphological evidence. *Canadian Journal of Zoology* **74**, 1045–1056.

Drevecky, C. (2011). Genetic differentiation of a sympatric lake-resident/anadromous species pair of threespine stickleback in Mud Lake, Alaska. M.S. Thesis, Department of Biological Sciences, DePaul University, Chicago, IL.

Evanno, G., Regnaut, S. & Goudet, J. (2005). Detecting the number of clusters of individuals using the software STRUCTURE: a simulation study. *Molecular Ecology* **14**, 2611-2620.

Falush, D., Stephens, M. & Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* **164**, 1567-1587.

Gelmond, O., von Hippel, F. A. & Christy, M. S. (2009). Rapid ecological speciation in three-spined stickleback Gasterosteus aculeatus from Middleton Island, Alaska: the roles of selection and geographic isolation. *Journal of Fish Biology* **75**, 2037-2051.

Griffiths, R., Orr, K. L., Adam, A. & Barber, I. (2000). DNA sex identification in the three-spined stickleback. *Journal of Fish Biology* **57,** 1331-1334.

Hellenthal, G., Busby, G. B., Band, G., Wilson, J. F., Capelli, C., Falush, D. & Myers, S. (2014). A genetic atlas of human admixture history. *Science* **343**, 747-751.

Hendry, A. P., Kaeuffer, R., Crispo, E., Peichel, C. L. & Bolnick, D. I. (2013). Evolutionary inferences from the analysis of exchangeability. *Evolution* **67**, 3429-3441.

Hendry, A. P. & Taylor, E. B. (2004). How much of the variation in adaptive divergence can be explained by gene flow: an evaluation using lake/stream stickleback. *Evolution* **58,** 2319-2331.

Hendry, A. P., Taylor, E. B., & McPhail, J. D. (2002). Adaptive divergence and the balance between selection and gen flow: lake and stream stickleback in the Misty system. *Evolution* **56,** 1199-1216.

Higuchi, M. & Goto, A. (1996) Genetic evidence supporting the existence of two distinct species in the genus *Gasterosteus* around Japan. *Environmental Biology of Fishes* **47**, 1–16.

Hoekstra, H., E., Krenz, J. G., & Nachman, M. W. (2005). Local adaptation in the rock pocket mouse (*Chaetodipus intermedius*): natural selection and phylogenetic history of populations. *Heredity* **94,** 217-228.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. & Cresko, W. A. (2010). Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**, e1000862-e1000862.

Hubisz, M. J., Falush, D., Stephens, M. & Pritchard, J. K. (2009). Inferring weak population structure with the assistance of sample group information. *Molecular Ecology Resources* **9**, 1322-1332.

Johnson, L.S. & Taylor, E.B. (2004) The distribution of divergent mitochondrial DNA lineages of threespine stickleback (*Gasterosteus aculeatus*) in the northeastern Pacific Basin: post-glacial dispersal and lake accessibility. *Journal of Biogeography* **31**, 1073–1083.

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J., Swofford, R., Pirun, M., Zody, M. C., White, S., Birney, E., Searle, S., Schmutz, J., Grimwood, J., Dickson, M. C., Myers, R. M., Miller, C. T., Summers, B. R., Knecht, A. K., Brady, S. D., Zhang, H., Pollen, A. A., Howes, T., Amemiya, C., Team, B. I. G. S. P. W. G. A., Baldwin, J., Bloom, T., Jaffe, D. B., Nicol, R., Wilkinson, J., Lander, E. S., Di Palma, F., Lindblad-Toh, K. & Kingsley, D. M. (2012). The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**, 55-61.

Klingenberg, C. P. (2011). MorphoJ: an integrated software package for geometric morphometrics. *Molecular Ecology Resources* **11,** 353-357.

Leinonen, T., Cano, J. M., Makinen, H. & Merila, J. (2006). Contrasting patterns of body shape and neutral genetic divergence in marine and lake populations of threespine sticklebacks. *Journal of Evolutionary Biology* **19**, 1803-1812.

Lescak, E.A., Marcotte, R.W., Kenney, L.A., von Hippel, F.A., Cresko, W.A., Sherbick, M.L., Colgren, J.J, and Lopez, J.A. 2014. Admixture of ancient mitochondrial lineages in three-spined stickleback populations from the North Pacific. *Journal of Biogeography* **42**, 532-539.

Liu, M., Lin, L., Gao, T., Yanagimoto, T., Sakurai, Y., & Grant, W. S. (2012). What maintains the central north Pacific genetic discontinuity in Pacific Herring? *PLoS One* **7,** e50340.

McGuigan, K., Nishimura, N., Currey, M., Hurwit, D., & Cresko, W. A. (2010). Cryptic genetic variation and body size evolution in threespine stickleback. *Evolution* **65,** 1203-1211.

Meirmans, P. & Van Tiendren, P. H. (2004). GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes* **4**, 792-794.

Moodie, G. E. E. & Reimchen, T. E. (1976). Phenetic variation and habitat differences in *Gasterosteus* populations of the Queen Charlotte Islands. *Systematic Zoology* **25**, 49-61.

Nielsen, R. & Wakeley, J. (2001). Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics* **158**, 885-896.

O'Reilly, P., Reimchen, T.E., Beech, R. & Strobeck, C. (1993) Mitochondrial DNA in *Gasterosteus* and Pleistocene glacial refugium on the Queen Charlotte Islands, British Columbia. *Evolution* **47**, 678–684.

Ortí, G., Bell, M.A., Reimchen, T.E. & Meyer, A. (1994) Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. *Evolution* **48**, 608–622.

Peichel, C. L., Nereng, K. S., Ohgi, K. A., Cole, B. L., Colosimo, P. F., Buerkle, C. A., Schluter, D. & Kingsley, D. M. (2001). The genetic architecture of divergence between threespine stickleback species. *Nature* **414**, 901-905.

Pritchard, J. K., Stephens, M. & Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* **155**, 945-959.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Ravinet, M., Prodohl, P. A., & Harrod, C. (2013). On Irish sticklebacks: morphological diversification in a secondary contact zone. *Evolutionary Ecology Research* **15**, 271-294.

Reimchen, T. E. (1988). Inefficient predators and prey injuries in a population of giant stickleback. *Canadian Journal of Zoology* **66**, 2036-2044.

Reimchen, T. E. (1994). Predators and evolution in threespine stickleback. In *Evolution of the Threespine Stickleback* (M. A. Bell & S. A. Foster, eds.), pp. 240-273. Oxford University Press.

Reist, J. D. (1986). An empirical evaluation of coefficients used in residual and allometric adjustment of size variation. *Canadian Journal of Zoology* **64,** 1363-1368.

Rohlf, F. J. (2010). TPSdig. Version 2.16.

Scribner, K. T, Page, K. S., & Bartron, M. L. (2001). Hybridization in freshwater fishes: a review of case studies and cytonuclear methods of biological inference. *Reviews in Fish Biology and Fisheries* **10**, 293-323.

Shapiro, M. D., Marks, M. E., Peichel, C. L., Blackman, B. K., Nereng, K. S., Jónsson, B., Schluter, D. & Kingsley, D. M. (2004). Genetic and developmental basis of evolutionary pelvic reduction in threespine sticklebacks. *Nature* **428**, 717-723.

Stacklies, W., Redestig, H., Scholz, M., Walther, D. & Selbig, J. 2007. pcaMethods—a bioconductor package providing PCA methods for incomplete data. *Bioinformatics* **23**, 1164-1167.

Vernot, B. & Akey, J. M. (2014). Resurrecting surviving Neandertal lineages from modern human genomes. *Science* **343**, 1017-1021.

Weigner, H. L. (2012) *Freshwater fish biogeography in the Bering Glacier region, Alaska*. PhD Thesis, University of Alaska Fairbanks, Fairbanks, AK.

Wu, T. D. & Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* **21,** 1859-1875.

Zakas, C., Jones, K., & Wares, J. P. (2014). Homogeneous nuclear background for mitochondrial cline in northern range of *Notochthamalus scabrosus*. *G3* **4,** 225-230.

Zelditch, M. L., Swiderski, D. L., Sheets, H. D., & Fink, W. L. (2004). *Geometric Morphometrics for Biologists*. San Diego: Elsevier.

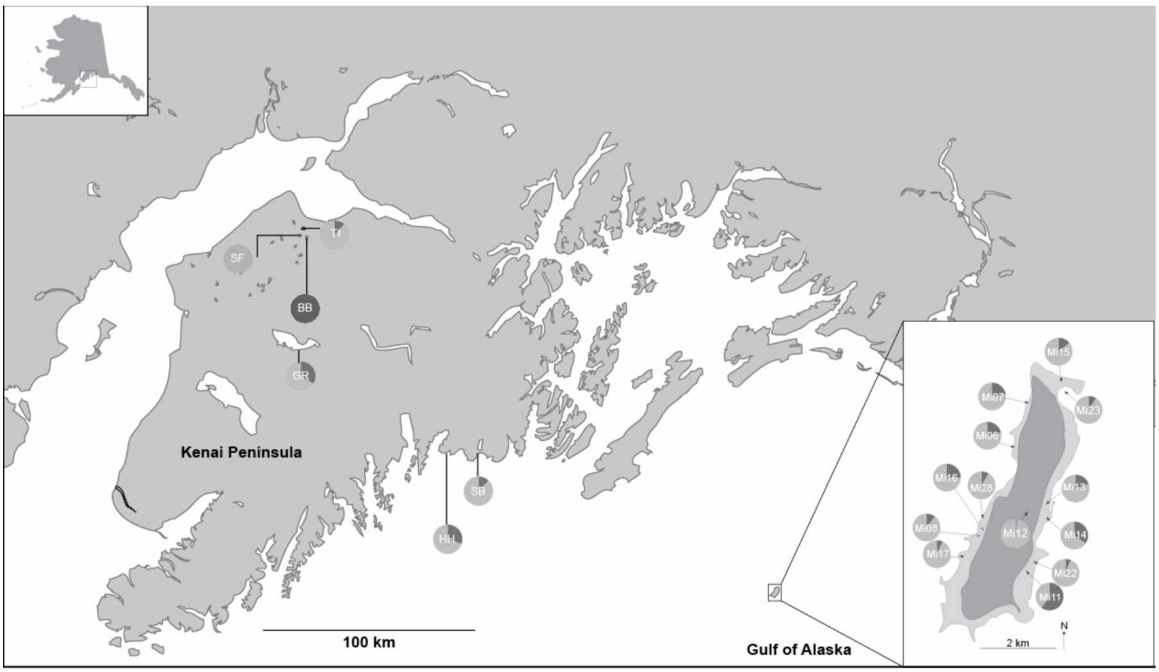

Figure 3.1. Sampling sites and clade proportions. Inset: Map of Alaska with box indicating sampling region. TI = Two Island, SF = Sportfish, BB = Barabara, GR = Grant, HH = Horsehead, SB = Salmonberry. Shading on Middleton Island represents its approximate pre-1964 size. Pie charts indicate clade proportions, with black representing proportion of the Trans North Pacific Clade and gray representing the Euro North American Clade.

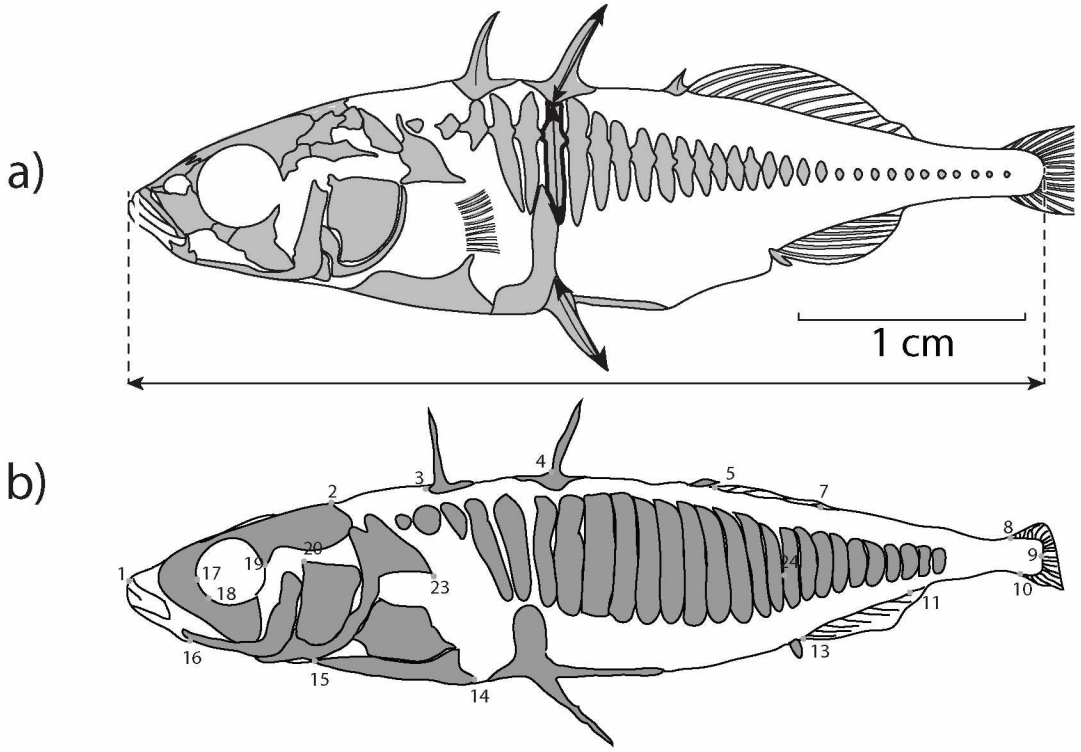

Figure 3.2. a) Armor measurements and b) Locations of landmarks for geometric morphometric analysis. Landmarks 1-23 are from McGuigan et al. (2010). Landmark 24 was added to unbend the fish.

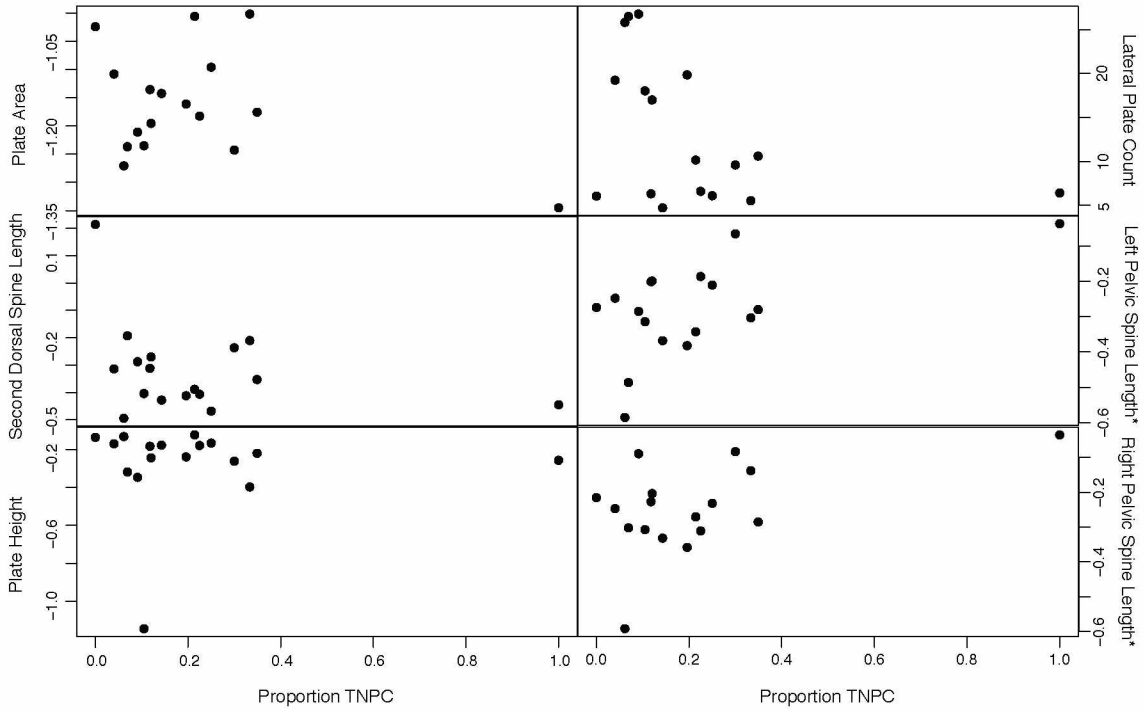

Figure 3.3. Mean trait values for each population plotted against proportion of individuals belonging to the Trans North Pacific Clade (TNPC). Both left and right pelvic spine lengths have significant positive correlations with TNPC proportion when all populations are analyzed. These relationships are no longer significant when Barabara is excluded.

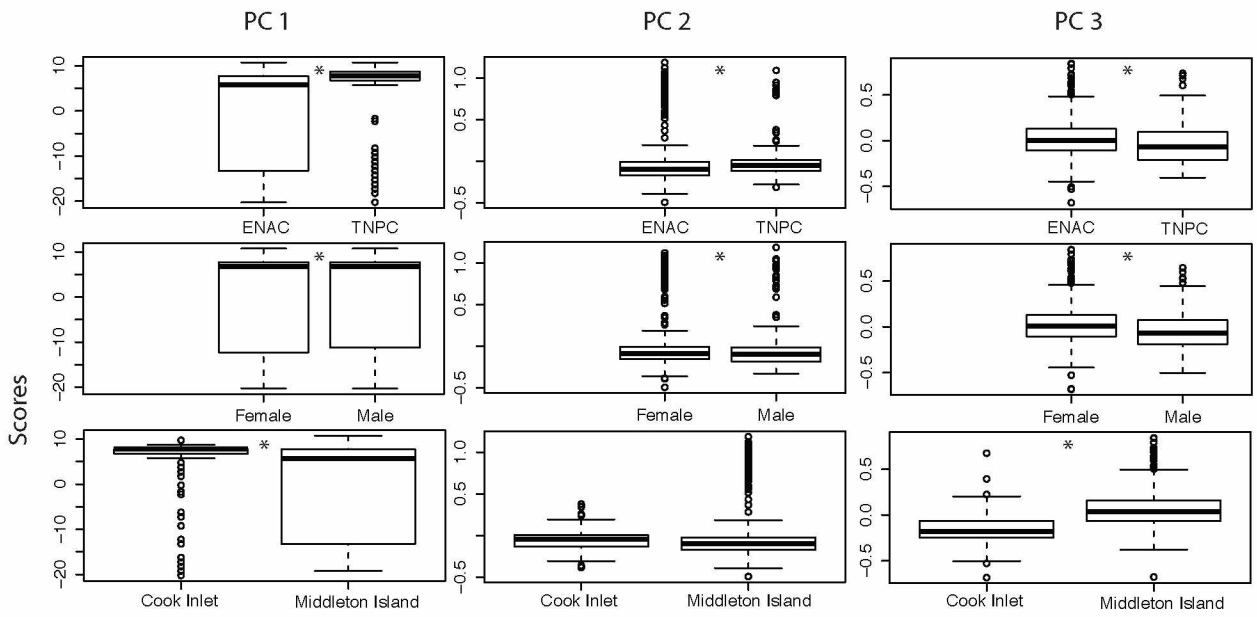

Figure 3.4. Comparisons of principal component (PC) scores in the analysis of armor phenotypes between clades, sexes, and regions. Asterisks indicate significant differences at $p<0.05$.

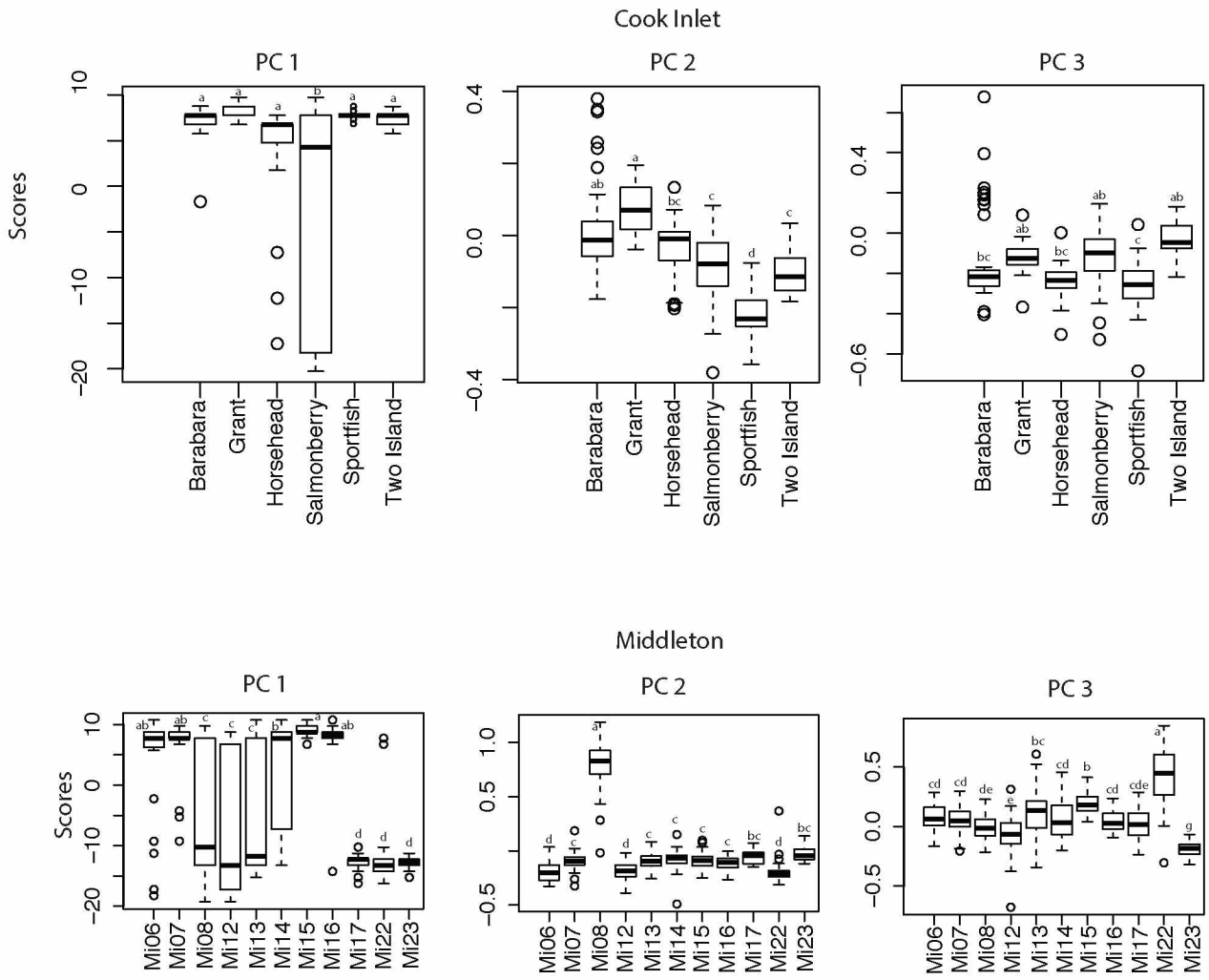

Figure 3.5. Comparisons of principal component (PC) scores in the analysis of armor phenotypes among populations from Cook Inlet (top) and Middleton Island (bottom). Letters indicate which sites significantly differ.

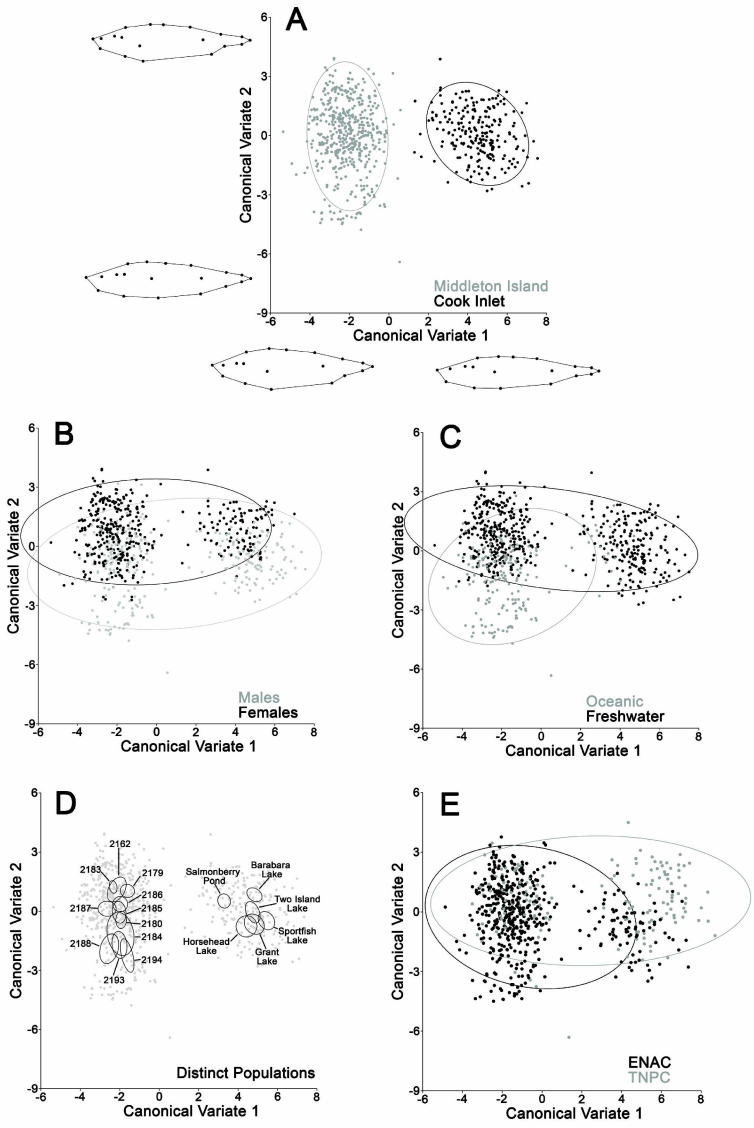

Figure 3.6. Plots of CV1 versus CV2 from two canonical variates analyses. Panels A-D depict results from a single analysis in which CVs were computed to distinguish between regions (Middleton Island versus Cook Inlet), sexes, ecotypes and among populations. Colors in each panel highlight different groupings within the same analysis: A) regions, B) sexes, C) ecotypes, D) populations. In A-C, ellipses are equal frequency ellipses containing 90% of the observations in respective groups, whereas ellipses in panel D are 95% confidence ellipses of population means. Wireframe diagrams in panel A represent shapes of fish at CV values of +/- 6.0. Panel E represents a similar analysis as that represented in A-D, but distinguished between clades in addition to regions, sexes and ecotypes. As such, the two populations for which clade identity is unknown were omitted (Two Island and Sportfish Lakes). In panel E, shape variation along each axis was similar to that for the analysis represented in panels A-D, and thus wireframe diagrams are not shown.

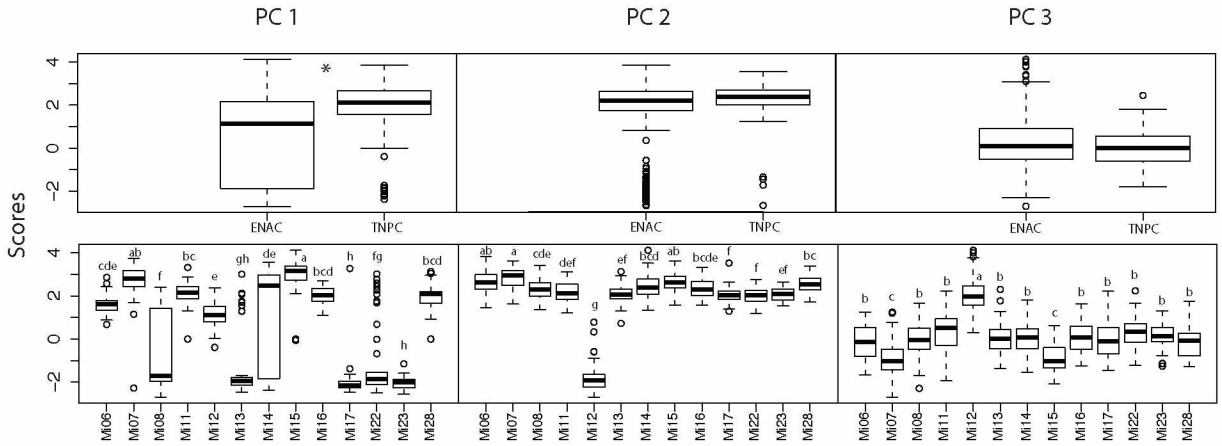

Figure 3.7. Comparisons of principal component (PC) scores from analysis of 1000 nuclear loci between clades (top) and among populations on Middleton Island (bottom). The asterisk indicates a significant difference at $p<0.05$. Letters represent sites that significantly differ.

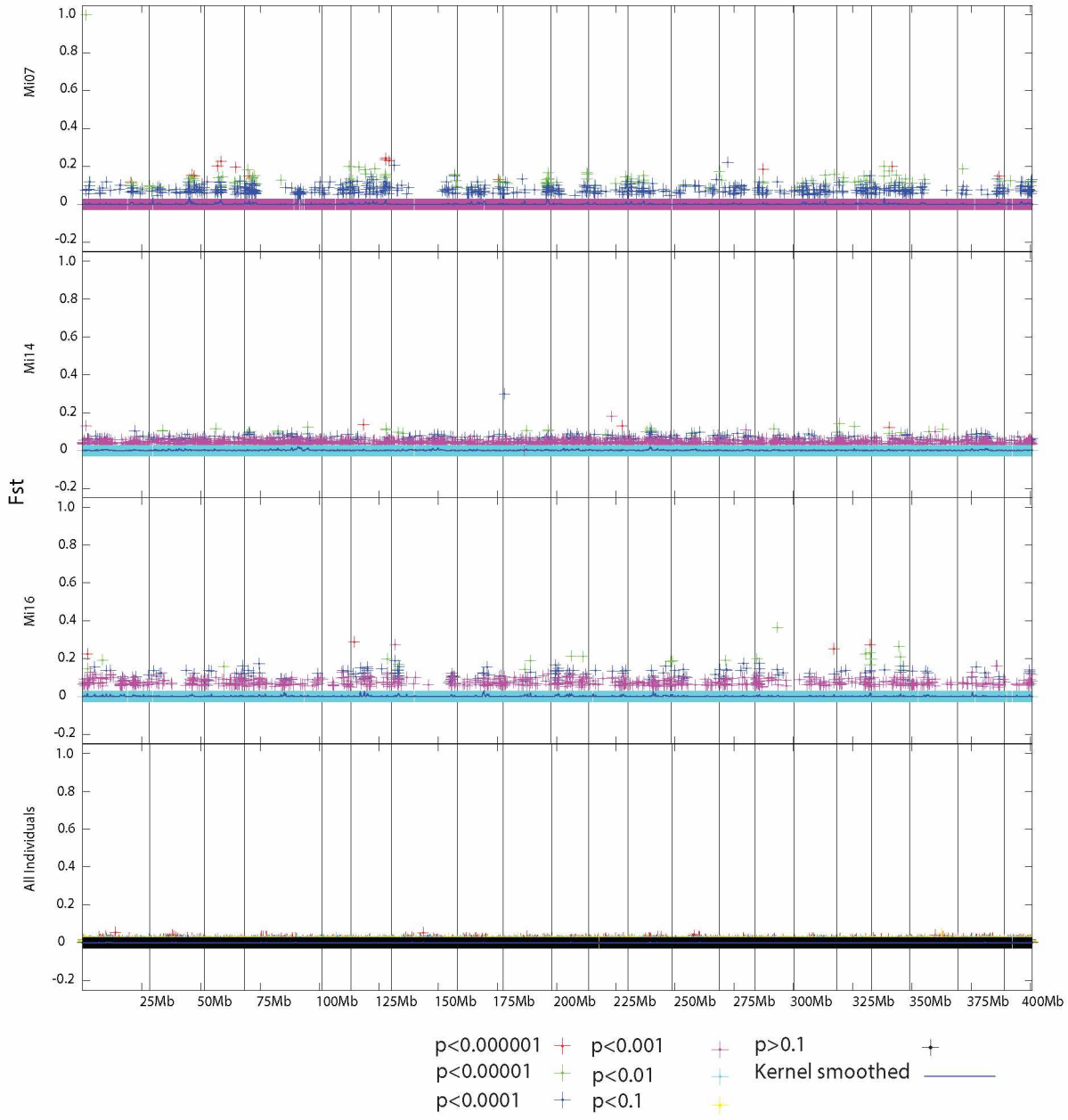

Figure 3.8. Plots of F_{ST} across the genome in comparisons of ENAC and TNPC individuals from Mi7, Mi14, Mi16, and all sites pooled do not reveal any nuclear loci that are strongly divergent between mitochondrial clades.

Table 3.1. Loadings for principal components (PCs) 1-3 in the analysis of armor phenotypes.

Trait	PC 1	PC 2	PC 3
Lateral Plate Count	-1.000	-0.005	-0.005
Plate Area	0.002	-0.171	-0.207
Plate Height	0.004	-0.980	0.096
Second Dorsal Spine Length	0.001	-0.097	-0.463
Left Pelvic Spine Length	0.005	-0.013	-0.590
Right Pelvic Spine Length	0.003	-0.009	-0.620

Table 3.2. Means of morphological measurements for individuals from each population (BB = Barabara, GR = Grant, HH = Horsehead, SB = Salmonberry, SF = Sportfish, TI = Two Island, Mi = Middleton). LP: left side lateral plate count, SDS: second dorsal spine length (mm), SPh: second supporting plate height (mm), SPa: second supporting plate area (mm^2), PS: pelvic score, LPS: left pelvic spine length (mm), RPS: right pelvic spine length (mm), SL: standard length (mm).

Population	n	LP	SDS	SPh	SPa	PS	LPS	RPS	SL
BB	59	6.29	4.20	6.89	0.70	5.44	0.11	0.12	65.1
GR	30	5.43	3.62	4.04	0.32	8.00	4.89	4.93	42.1
HH	30	9.50	5.30	5.05	0.53	8.00	7.87	7.66	50.7
SB	50	16.9	3.93	4.78	0.47	8.00	5.66	5.63	47.3
SF	29	5.97	3.40	4.65	0.39	8.00	4.10	4.06	46.9
TI	30	6.23	3.95	5.28	0.49	8.00	4.91	4.85	50.6
Mi06	30	6.23	3.95	5.28	0.49	8.00	4.91	4.85	50.6
Mi07	30	6.23	3.95	5.28	0.49	8.00	4.91	4.85	50.6
Mi08	86	18	5.20	8.27	0.62	8.00	6.89	7.06	64.0
Mi12	99	19.2	4.53	6.80	0.51	8.00	5.18	5.14	51.1
Mi13	46	19.8	6.48	7.88	1.35	8.00	7.63	8.02	70.4
Mi14	63	10.5	4.74	6.19	0.76	8.00	5.66	5.66	55.6
Mi15	42	4.62	3.77	5.48	0.55	8.00	4.34	4.71	50.2
Mi16	44	6.00	3.43	4.86	0.43	8.00	4.20	4.22	42.8
Mi17	29	26.41	6.54	7.84	1.29	8.00	8.45	8.76	70.5
Mi22	49	25.73	6.23	7.50	1.13	8.00	7.90	7.84	66.3
Mi23	22	26.68	6.37	6.99	1.09	8.00	8.20	8.25	66.6

Table 3.3. Pearson correlations between each trait mean and proportion Trans North Pacific clade.

Trait	Full Dataset		Without Barabara	
Lateral plate count	t=-1.74	p=0.102	t=-2.07	p=0.057
Right pelvic spine length	t=2.20	p=0.044	t=1.07	p=0.303
Left pelvic spine length	t=2.64	p=0.019	t=1.49	p=0.157
Second dorsal spine length	t=-1.27	p=0.223	t=-1.19	p=0.255
Plate height	t=0.16	p=0.879	t=0.174	p=0.865
Plate area	t=-1.57	p=0.138	t=0.632	p=0.538

Table 3.4. Univariate tests of significance, effect sizes, and powers for canonical variates (CV) 1 and 2 for the entire dataset and the analysis excluding Sportfish (SF) and Two Island (TI).

ALL POPS

	CV1				
Effect	F/R	SS	df	F	p
Intercept	Fixed	412.00	1	77.34	0.000000
Region	Fixed	4058.30	1	647.50	0.000000
Pop(Region)	Random	137.01	15	10.58	0.000000
Ecotype	Fixed	0.33	1	0.38	0.536792
Sex	Fixed	25.41	1	29.41	0.000000
Error		576.11	667		
	CV2				
Effect	F/R	SS	df	F	p
Intercept	Fixed	106.34	1	17.23	0.000645
Region	Fixed	0.11	1	0.02	0.902841
Stock(Region)	Random	159.59	15	11.41	0.000000
Ecotype	Fixed	101.24	1	108.57	0.000000
Sex	Fixed	381.16	1	408.75	0.000000
Error		621.98	667		

NO TI/SF

	CV1				
Effect	F/R	SS	df	F	p
Intercept	Fixed	495.17	1	82.23	0.000000
Region	Fixed	3328.56	1	423.39	0.968443
Pop(Region)	Random	133.76	13	11.56	0.197731
Ecotype	Fixed	1.34	1	1.51	0.219987
Sex	Fixed	20.46	1	22.96	0.000002
Clade	Fixed	2.48	1	2.79	0.095656
Error		542.81	609		
	CV2				
Effect	F/R	SS	df	F	p
Intercept	Fixed	91.62	1	10.93	0.005033
Region	Fixed	15.60	1	1.41	0.255324
Pop(Region)	Random	189.74	13	15.99	0.000000
Ecotype	Fixed	108.40	1	118.77	0.000000
Sex	Fixed	284.04	1	311.22	0.000000
Error		555.82	609		

Table 3.5. Summary genetic statistics for all sites calculated for only nucleotide positions that are polymorphic in at least one site ("Variant") and all nucleotide positions across all RAD sites regardless of whether they are polymorphic or fixed ("All"). These statistics include the average number of individuals genotyped per locus (N), the number of variable sites unique to each site (Private), the number of sites (Sites), percentage of polymorphic loci (% Polymorphic), the average frequency of the major allele (P), the average observed and expected homozygosity and heterozygosity per locus, the average nucleotide diversity (π), and Wright's inbreeding coefficient (F_{IS}).

Variant Positions									
	Private	N	P	Obs. Het.	Obs. Hom.	Exp. Het.	Exp. Hom.	π	F_{IS}
TNPC	6020	131.29	0.95	0.05	0.95	0.07	0.93	0.07	0.08
ENAC	40037	675.22	0.95	0.05	0.95	0.07	0.93	0.07	0.10
All Positions (Variant and Fixed)									
	Sites	% Polymorphic	P	Obs. Het.	Obs. Hom.	Exp. Het.	Exp. Hom.	π	F_{IS}
TNPC	2971848	0.05	1.00	0.00	1.00	0.00	1.00	0.00	0.00
ENAC	2970239	0.05	1.00	0.00	1.00	0.00	1.00	0.00	0.00

Table 3.6. Summary genetic statistics for all sites calculated for only nucleotide positions that are polymorphic in at least one site ("Variant") and all nucleotide positions across all RAD sites regardless of whether they are polymorphic or fixed ("All") in TNPC and ENAC individuals from Mi07, Mi14, and Mi16. These statistics include the average number of individuals genotyped per locus (N), the number of variable sites unique to each site (Private), the number of sites (Sites), percentage of polymorphic loci (% Polymorphic), the average frequency of the major allele (P), the average observed and expected homozygosity and heterozygosity per locus, the average nucleotide diversity (π), and Wright's inbreeding coefficient (F_{IS}).

Variant Positions									
	Private	n	P	Obs. Het.	Obs. Hom.	Exp. Het.	Exp. Hom.	π	F_{IS}
Mi07 TNPC	506	9.60	0.93	0.09	0.91	0.09	0.90	0.10	0.02
Mi07 ENAC	2075	31.49	0.93	0.09	0.91	0.10	0.90	0.10	0.03
Mi14 TNPC	4366	25.86	0.91	0.11	0.89	0.13	0.87	0.13	0.06
Mi14 ENAC	8587	37.46	0.91	0.11	0.89	0.13	0.87	0.13	0.07
Mi16 TNPC	982	10.78	0.92	0.11	0.89	0.12	0.88	0.12	0.02
Mi16 ENAC	3226	28.44	0.92	0.12	0.89	0.12	0.88	0.12	0.03
All Positions									
	Sites	Variant Sites	P	Obs. Het.	Obs. Hom.	Exp. Het.	Exp. Hom.	π	F_{IS}
Mi07 TNPC	3014838	81389	1.00	0.00	1.00	0.00	1.00	0.00	0.00
Mi07 ENAC	3014836	81242	1.00	0.00	1.00	0.00	1.00	0.00	0.00
Mi14 TNPC	3014704	81114	1.00	0.00	1.00	0.00	1.00	0.00	0.00
Mi14 ENAC	3014555	80959	1.00	0.00	1.00	0.00	1.00	0.00	0.00
Mi16 TNPC	3014853	81286	1.00	0.00	1.00	0.00	1.00	0.00	0.00
Mi16 ENAC	3014721	81140	1.00	0.00	1.00	0.00	1.00	0.00	0.00

Table 3.7. Heat map of pairwise F_{ST} in TNPC and ENAC individuals from Mi07, Mi14, and Mi16. Green represents least divergent comparisons and red represents the greatest pairwise difference.

	Mi07 ENAC	Mi14 TNPC	Mi14 ENAC	Mi16 TNPC	Mi16 ENAC
Mi07 TNPC	0.00	0.01	0.02	0.03	0.03
Mi07 ENAC		0.03	0.02	0.07	0.04
Mi14 TNPC			0.00	0.01	0.01
Mi14 ENAC				0.02	0.01
Mi16 TNPC					0.00

Conclusion

This dissertation describes a 'natural experiment' for studying contemporary evolution in wild populations. We report the presence of resident freshwater threespine stickleback on previously submarine terraces that were uplifted in 1964. We demonstrate using a dense genomic dataset that oceanic stickleback were the most probable colonizers of new freshwater habitats and that they are continuously invading these ponds.

Despite secondary contact, resident freshwater stickleback diverged from their oceanic ancestors in a suite of phenotypes associated with armor, body size, foraging, and swimming. Many of the traits examined have known genetic bases (Peichel *et al.*, 2001, Cresko *et al.*, 2004; Colosimo *et al.*, 2004, 2005, Chan *et al.*, 2010), suggesting that, to a large degree, the phenotypic variation we observe between ecotypes is due to genetic divergence rather than trait plasticity. Furthermore, the measured genetic divergence between ecotypes is nearly the same in magnitude as that documented in populations that are tens of thousands of years old (Hohenlohe *et al.*, 2010), supporting the hypothesis that much of the evolution of freshwater phenotypes occurs shortly after colonization.

The variation that we report does not seem to be influenced by an ancient period of allopatry that led to deep divergence in the mitochondrial genome – divergence that we can still readily observe in extant populations. We propose that this ancient signature has been erased in the nuclear genome due to recombination and a prolonged period of secondary contact between the two ancient lineages throughout the North Pacific. The presence of a 6,000 km zone of admixture between mitochondrial lineages in this region (Lescak *et al.*, 2014) and absence of nuclear genetic structure between lineages suggests that this region is a corridor for the widespread movement of stickleback.

Much attention has been paid recently to answering questions about how the structure of the stickleback species complex and organization of its genome allow for rapid, parallel evolution of stereotypical freshwater phenotypes. The 'transporter hypothesis' (Schluter & Conte, 2009) proposes that low levels of hybridization between oceanic (most likely anadromous) and resident freshwater stickleback maintain freshwater alleles at low frequency in the oceanic stock.

Supporting the transporter hypothesis is the finding that the same genes, and often the same alleles, underlie phenotypic variation in geographically disparate populations. The role of the *Ectodysplasin* locus has been implicated in lateral plate reduction in stickleback throughout their distribution (Cresko *et al.*, 2004; Colosimo *et al.*, 2004, 2005; Hohenlohe *et al.*, 2010; Jones *et al.*, 2012). Furthermore, Hohenlohe *et al.* (2010) and Jones *et al.* (2012) were able to demonstrate that patterns of genomic divergence between oceanic and resident freshwater ecotypes are remarkably similar in independent populations. These studies support the important role of hybridization in maintaining freshwater genotypes at low frequency in oceanic populations, which allows for parallel genomic evolution upon invasion of freshwater environments.

How can a suite of complex phenotypes evolve so rapidly? Pleiotropy, or one genomic region underlying multiple phenotypes, has been identified in loci underlying lateral plate width and height (Colosimo *et al.*, 2004) as well as body shape and lateral plate number (Peichel *et al.*, 2001; Jones *et al.*, 2012), which means that a change to any of those genomic regions will influence multiple phenotypes. Chromosomal inversions, which suppress recombination and allow combinations of freshwater alleles to remain linked despite secondary contact with oceanic stickleback, overlap regions of the genome that

influence multiple ecologically important phenotypes (Jones *et al.*, 2012). Reduced recombination has been detected in additional genomic regions under strong selection (Hohenlohe *et al.*, 2012), which may be due to the presence of inversions that have yet to be identified.

Linkage disequilibrium also maintains associations between loci on different linkage groups that underlie ecologically important phenotypes. *Ectodysplasin*, which controls lateral plate development, has been found to be in association with *Pitx1*, which plays a role in pelvic girdle development (Chan *et al.*, 2010; Hohenlohe *et al.*, 2012). This feature of the stickleback's genomic architecture could explain the repeated occurrence of reductions in lateral plate number and pelvic girdle elements across resident freshwater populations (e.g., Bell & Orti, 1994).

Standing genetic variation in the stickleback's genome allows oceanic colonizers to adapt to freshwater environments within just a few decades (or perhaps just a few generations; Barrett, 2010). Our work provides the much-needed evidence that contemporary change in wild populations can be due to evolution, rather than plasticity without associated genetic change. Species potential for contemporary evolution has important implications for human-induced environmental change and the conservation of biodiversity; the degree to which different species of conservation concern have the ability to adapt to abrupt environmental changes is an important question for future research.

Where do we go from here? The biomedical field has largely advanced our understanding of intricate phenotype-genotype relationships. Stickleback are poised to be a major model for biomedical research focused on identifying genomic regions – and even alleles – underlying phenotypes of developmental interest. Much of the stickleback's

genome is conserved across vertebrates (Jeffrey, 2005; Tian & Price, 2005; Protas *et al.,* 2006, 2007), which allows researchers to make connections to human health. With the massive phenotypic and genetic datasets we have collected, we have the potential to apply genome-wide association techniques to uncover the genetic basis for a variety of traits of interest and investigate both epistatic and pleiotropic relationships among genes underlying ecologically- and developmentally-important phenotypes.

Future work can also improve our understanding of cytonuclear disequilibrium, or non-random associations between mitochondrial and nuclear genes. These relationships have previously been detected across taxa (e.g. Scribner *et al.*, 1999; Sambatti *et al.*, 2008; Barr & Fishman, 2010). While we did not detect them in our study, future investigations can test specifically for relationships between mitochondrial-encoded genes in the nuclear genome and variation in mtDNA.

This work has provided a rare example of contemporary evolution in the wild in which populations have a known approximate age and the ancestral population can be identified. Our work supports the 'transporter hypothesis' that the stickleback's genome is rapidly reassembled when oceanic individuals invade freshwater habitats. We propose that the bulk of evolution in freshwater stickleback occurs over decadal timescales immediately after colonization, rather than over thousands of years, and challenge the idea that the evolution of complex traits is necessarily a slow and gradual process.

References

Barr, C. M. & Fishman, L. 2010. The nuclear component of a cytonuclear hybrid incompatibility in Mimulus maps to a cluster of pentatricopeptide repeat genes. *Genetics* **184**: 455-465.

Barrett, R. D. H. 2010. Adaptive evolution of lateral plates in three-spined stickleback *Gasterosteus aculeatus*: a case study in functional analysis of natural variation. *J Fish Biol* **77**: 311-328.

Bell, M. A. & Orti, G. 1994. Pelvic reduction in threespine stickleback from Cook Inlet lakes: geographic distribution and intrapopulation variation. *Copeia* **1994**: 314-325.

Chan, Y. F., Marks, M. E., Jones, F. C., Villarreal, G., Shapiro, M. D., Brady, S. D. et al. 2010. Adaptive evolution of pelvic reduction in sticklebacks by recurrent deletion of a *Pitx1* enhancer. *Science* **327**: 302-305.

Colosimo, P. F., Hosemann, K. E., Balabhadra, S., Villarreal, G., Dickson, M., Grimwood, J. et al. 2005. Widespread parallel evolution in sticklebacks by repeated fixation of Ectodysplasin alleles. *Science* **307**: 1928-1933.

Colosimo, P. F., Peichel, C. L., Nereng, K., Blackman, B. K., Shapiro, M. D., Schluter, D. et al. 2004. The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology* **2**: E109.

Cresko, W. A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P. et al. 2004. Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proc Natl Acad Sci USA* **101**: 6050-6055.

Hohenlohe, P. A., Bassham, S., Currey, M. & Cresko, W. A. 2012. Extensive linkage disequilibrium and parallel adaptive divergence across threespine stickleback

genomes. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **367**: 395-408.

Hohenlohe, P. A., Bassham, S., Etter, P. D., Stiffler, N., Johnson, E. A. & Cresko, W. A. 2010. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genetics* **6**: e1000862-e1000862.

Jeffrey, W. R. 2005. Adaptive evolution of eye degeneration in the Mexican blind cavefish. *J Hered* **96**: 185-196.

Jones, F. C., Grabherr, M. G., Chan, Y. F., Russell, P., Mauceli, E., Johnson, J. et al. 2012. The genomic basis of adaptive evolution in threespine sticklebacks. *Nature* **484**: 55-61.

Lescak, E. A., Marcotte, R. W., Kenney, L. A., Hippel, F. A., Cresko, W. A., Sherbick, M. L. et al. 2014. Admixture of ancient mitochondrial lineages in three spined stickleback populations from the North Pacific. *J Biogeogr* **42**: 532-539.

Peichel, C. L., Nereng, K. S., Ohgi, K. A., Cole, B. L., Colosimo, P. F., Buerkle, C. A. et al. 2001. The genetic architecture of divergence between threespine stickleback species. *Nature* **414**: 901-905.

Protas, M., Conrad, M., Gross, J. B., Tabin, C. & Borowsky, R. 2007. Regressive evolution in the Mexican cave tetra, Astyanax mexicanus. *Curr Biol* **17**: 452-454.

Protas, M. E., Hersey, C., Kochanek, D., Zhou, Y., Wilkens, H., Jeffery, W. R. et al. 2006. Genetic analysis of cavefish reveals molecular convergence in the evolution of albinism. *Nat Genet* **38**: 107-111.

Sambatti, J. B., Ortiz-Barrientos, D., Baack, E. J. & Rieseberg, L. H. 2008. Ecological selection maintains cytonuclear incompatibilities in hybridizing sunflowers. *Ecol Lett* **11**: 1082-1091.

Schluter, D. & Conte, G. L. 2009. Genetics and ecological speciation. *Proceedings of the National Academy of Sciences* **106 Suppl 1**: 9955-9962.

Scribner, K. T., Datta, S., Arnold, J. & Avise, J. C. 1999. Empirical evaluation of cytonuclear models incorporating genetic drift and tests for neutrality of mtDNA variants: data from experimental Gambusia hybrid zones. *Genetica* **105**: 101-108.

Tian, N. M. & Price, D. J. 2005. Why cavefish are blind. *BioEssays* **27**: 235-238.

Appendix: Admixture of Ancient Mitochondrial Lineages in Three-Spined Stickleback Populations from the North Pacific[1]

Abstract

Aim We test the hypothesis that the North Pacific coastline from British Columbia to the Kuril Islands forms a broad region of admixture between two divergent mitochondrial-genome (mtDNA) lineages in three-spined stickleback (*Gasterosteus aculeatus*): the Euro–North American (ENA) and trans-North-Pacific (TNP or Japanese) clades. We test whether distance is the primary determinant of geographical patterns of haplotype distributions and whether deep-water trenches in the Aleutian and Kuril archipelagos impede gene flow.

Location Coastal marine and freshwater sites from the Kuril Islands (north-western Pacific Ocean) to Oregon (north-eastern Pacific Ocean).

Methods We determined the mtDNA clade for 1327 individuals from 67 locations across 8000 km of the North Pacific using restriction fragment length polymorphism assays of the cytochrome *b* mitochondrial gene. We supplemented this with published clade designations from coastal Pacific populations and tested whether distance is the primary determinant of geographical patterns of haplotype distributions using a generalized linear model and Mantel test. We also used analysis of molecular variation (AMOVA) to test for significant partitioning of genetic variation by deep-water trenches.

Results The western boundary of the ENA clade was Simushir in the Kuril Islands and the eastern boundary of the TNP clade was British Columbia. Coastline distance from Japan was a significant predictor of TNP abundance. Clade composition variance was high at

---

[1] Lescak, E.A., Marcotte, R.W., Kenney, L.A., von Hippel, F.A., Cresko, W.A., Sherbick, M.L., Colgren, J.J, and Lopez, J.A. 2014. Journal of Biogeography 42: 532-539.

small geographical scales, in apparent discordance with the broader-scale pattern of admixture. Deep-water trenches were not found to significantly partition genetic variation in the Aleutian and Kuril Island chains.

Main conclusions The North Pacific coastline forms a broad region of secondary contact between divergent mitochondrial lineages, but with clear western and eastern boundaries. Patterns of clade abundance in the Kuril and Aleutian islands are similar to those observed in other taxa, suggesting shared biogeographical histories and common barriers shaping species distributions along the North Pacific coast. Overall patterns of clade distributions are likely to be driven by a combination of factors including geomorphological impediments to migration, ecology, asymmetrical dispersal patterns and differences in timing of population expansions.

Introduction

Identifying how landscape and the environment shape patterns of genetic diversity improves our understanding of microevolution (Storfer *et al.*, 2006). The distribution of genetic variability at neutral loci has been shown to be associated with geography (e.g. isolation by distance or ecological discontinuities) in marine and freshwater fish (Jørgensen *et al.*, 2005; Dionne *et al.*, 2008, Jones *et al.*, 2012). However, an open question concerns how well genetic markers are able to recapture deep ancestral divergence over large geographical scales.

A key organism to address this question is the three-spined stickleback (*Gasterosteus aculeatus* Linnaeus, 1758) because it is broadly distributed in temperate and subpolar freshwater and marine habitats throughout the Northern Hemisphere and many aspects of its biology have been well studied for decades (for reviews, see Bell & Foster,

1994; Östlund-Nilsson *et al.* 2006; von Hippel, 2010). A growing body of research on the adaptive evolution of three-spined stickleback to freshwater and marine environments reveals linkages between genetic variants and ecologically important phenotypes (e.g. Peichel *et al.*, 2001; Colosimo *et al.*, 2004, 2005; Cresko *et al.*, 2004; Jones *et al.*, 2012). To reconstruct the biogeographical history of this species and to test hypotheses regarding present and historical gene flow, we examined the genetic variation in populations from the North Pacific. This deep history allows us to more fully understand the evolutionary dynamics of adaptive radiation, by enabling the characterization of local, regional and global patterns of genetic variability and the spatial factors that influence the distribution of that variation (e.g. Jones *et al.*, 2012).

Analyses of both nuclear and mitochondrial (mtDNA) genomes reveal distinct lineages in modern populations of three-spined stickleback found in the Atlantic and Pacific basins (Haglund *et al.*, 1992; Ortí *et al.*, 1994; Colosimo *et al.*, 2005; Jones *et al.*, 2012) – the Euro–North American (ENA) and Japanese or trans-North-Pacific (TNP) clades – named based on assessments of their geographical distribution (Ortí *et al.*, 1994; Johnson & Taylor, 2004). All surveys of standing mtDNA genetic variation reported to date have confirmed that only these two lineages occur in extant freshwater and marine three-spined stickleback populations and each clade has multiple sub-lineages (Gach & Reimchen, 1989; Haglund *et al.*, 1992; O'Reilly *et al.*, 1993; Ortí *et al.*, 1994; Deagle *et al.*, 1996; Cresko, 2000; Mäkinen *et al.*, 2008; Weigner, 2012).

Divergence between the two mtDNA lineages has been estimated at approximately 0.9–1.3 Ma based on silent-site substitutions in cytochrome *b* DNA sequences, a rate of divergence calibrated from the split between nine-spined stickleback (*Pungitius pungitius*)

and three-spined stickleback, the fossil record and Markov chain Monte Carlo approaches to modelling migration rates (Ortí *et al.*, 1994; Nielsen & Wakeley, 2001). The TNP clade is most common in coastal areas of the western Pacific basin and is not found in any surveyed populations from the Atlantic Ocean, while the ENA clade is found in the eastern Pacific and in the Atlantic (O'Reilly *et al.*, 1993; Ortí *et al.*, 1994; Mäkinen *et al.*, 2008).

The first global survey of mtDNA diversity revealed that Alaska and British Columbia are regions of clade admixture (Ortí *et al.*, 1994). The observations reported to date point to British Columbia as the eastern extent of the TNP clade members (Ortí *et al.*, 1994; Deagle *et al.*, 1996), whereas the ENA lineage has not been reported further west than south-central Alaska. Clade-specific haplotype frequencies have been documented among populations in close proximity in the region of admixture; the proportion of individuals carrying ENA mtDNA haplotypes ranges from 0.03 to 1 in anadromous populations and from 0 to 1 in resident freshwater populations (Deagle *et al.*, 1996; Cresko, 2000; Weigner, 2012).

Surveys of mtDNA clade diversity in three-spined stickleback populations from the Pacific have left an unsampled gap, approximately 4600 km long, between south-central Alaska and Japan (Fig. A.1). On the western Pacific coast, the northernmost population sampled is on Hokkaido Island and the westernmost samples from North America originated from Cook Inlet, Alaska (Ortí *et al.*, 1994; Cresko, 2000). Until this report, the distribution and composition of mtDNA clades along the length of the Kuril and Aleutian island chains had not been examined, but we hypothesize that these are likely zones of secondary contact between the mtDNA lineages.

In this study, we test the hypothesis that the North Pacific coastline from British Columbia to the Kuril Islands forms a broad region of admixture between the ENA and TNP clades. We also test whether distance is the primary determinant of geographical patterns of haplotype distributions and if deep-water trenches in the Aleutian and Kuril archipelagos impede gene flow. We characterized the distribution of the two mtDNA clades across the North Pacific basin in populations along an approximately 8000 km transect from the Kuril Islands to Oregon (Fig. A.1). If the distance from source populations shapes clade composition, then we would expect to see a cline in clade proportions with increasing frequency of TNP haplotypes progressing westwards along the North Pacific Basin toward the source populations. We hypothesize that discontinuities in this cline may represent partial barriers to migration due to present or historical impediments to fish movement, such as deep-water trenches in the Kuril and Aleutian island chains, and/or ecotones.

Materials and Methods

Sampling

We determined clade assignment of the mtDNA genome from 1327 individuals originating from 67 sites across approximately 8000 km of the coastal North Pacific from the Kuril Islands to Oregon. Tissue samples from specimens from Washington, the Kuril Islands and Sakhalin were lent by the University of Washington Fish Collection, where the vouchers are archived. We supplemented this dataset with published mtDNA clade frequencies from Cook Inlet, Alaska (Cresko, 2000; E.A.L., unpublished data), Bering Glacier, Alaska (Weigner, 2012) and British Columbia (Deagle *et al.*, 1996, Johnson & Taylor, 2004), resulting in a total dataset of 3682 individuals from 169 locations (Fig. A.1 and Table A.1). Specimens were collected between 1994 and 2011 using minnow traps, beach seines and

dip nets. Fish and/or tissue samples were preserved in ethanol. DNA was isolated from fin clips using the Chelex method (Bio-Rad, Hercules, CA, USA) or either DNeasy or Puregene (Qiagen, Valencia, CA, USA) tissue kits following manufacturers' protocols.

Haplotype determination

The following primer pairs were used to amplify a region of the cytochrome *b* gene: Gacu_14459_F (5'-TGATGAAACTTTGGTTCCCTCCTT-3') and Gacu_15177_R (5'-TTGATGTGAGGTGGAGTGACTAA-3'), or 14372 (5'-ATGGCAAGCCTACGAAAAACGCAC-3') and 15100 (5'-TGCTAGGGATGTAAGGGCAATTAG-3'; Weigner, 2012). Amplification conditions were: 2 min at 94 °C, following by 28 cycles of 30 s at 94 °C, 15 s at 58 °C and 45 s at 72 °C, and finally 7 min at 72 °C, or 6 min at 94 °C, followed by 37 cycles of 60 s at 94 °C, 60 s at 63 °C and 120 s at 72 °C and finally 5 min at 72 °C, for each of the primer pairs respectively.

To assign the amplified cytochrome *b* gene fragments to either the TNP or the ENA clade, we performed restriction digests with *BstXI* and *NlaIII* and size-separated the products on agarose gels. The amplified cytochrome *b* fragment in the TNP clade lacks *BstXI* recognition sites and includes one *NlaIII* site, whereas descendants of the ENA lineage have one *BstXI* recognition site and two *NlaIII* sites (Ortí *et al.*, 1994). We performed independent digests with each of the two enzymes to increase confidence in clade assignment. All individuals were unambiguously assigned to one lineage based on the diagnostic restriction sites. We also sequenced the amplified gene fragment in a subset of 11 individuals (6 ENA and 5 TNP) from six Alaskan populations to confirm consistent correspondence between restriction-site arrangement and mtDNA clade affiliation.

Biogeographical analyses

To characterize relationships between geography and the distribution of the two mtDNA clades across populations, we first determined direct and coastline distances between each sampling location and Sakhalin, the westernmost sampled population, using GOOGLE EARTH and ARCGIS 10.1 (ESRI, Redlands, CA, USA). Because direct and coastline distances are highly correlated ($r^2 = 0.99$), we used coastline distances only in our analyses, because oceanic stickleback are thought to move primarily along the coast rather than through open water (Bell & Foster, 1994). In some instances, however, stickleback have been found in offshore regions of the Pacific Ocean (Quinn & Light, 1989; Deagle *et al.*, 1996; Morita *et al.*, 2009; Atcheson *et al.*, 2012a,b).

To whether distance is the primary determinant of geographical patterns of haplotype distributions, we first fitted a binomial generalized linear model to the relationship between coastline distance and clade proportions. This analysis was initially performed using all populations. We were able to categorize individuals from Alaska, British Columbia and Oregon into ecotypes (freshwater versus oceanic) based on phenotype (high versus low lateral plate counts and body size) or previously published classifications, which allowed us to also perform separate analyses on the two ecotypes from these regions. Outliers were identified through the use of diagnostic plots of residuals to determine which sampling sites did not follow the linear model. Secondly, we used a Mantel test to test for significance between pairwise F_{ST} and geographical distances among populations.

To test whether deep-water trenches impede migration, we tested for significant partitioning of genetic variation using an analysis of molecular variance (AMOVA; Excoffier

et al., 1992), grouping populations in the Kuril and Aleutian archipelagos by their locations on either side of the Bussol' Strait and Amchitka Pass. Because of our unequal sampling distribution across sites (Table A.1), we resampled 16 individuals (the mean number of individuals collected across populations) from all admixed sites 1000 times with replacement within each replicate. We then re-ran all analyses with the resampled distributions to ensure that our results were not due to unequal sample sizes. All analyses were performed in R 2.11.1 (R Core Team, 2014) or GENODIVE (Meirmans & van Tienderen, 2004).

Results

Overall clade proportions and clade distribution boundaries

Overall, 78% of the surveyed stickleback carried mtDNA from the ENA clade (Fig. A.1, Table A.1), which partly reflects a more intensive sampling of locations and individuals from the eastern Pacific coast. When we resampled our dataset, the ENA haplotype still predominated, but the percentage of individuals decreased to 58%. We found the western distribution limit of the ENA lineage to lie at the island of Simushir in the Kuril chain. The eastern distribution limit of the TNP clade lay along the coast of British Columbia (Johnson & Taylor, 2004; Fig. A.1). We found a great deal of heterogeneity in clade proportions among closely spaced populations (Figs A.1 & A.2), with the Aleutian chain, Kenai Peninsula and Bering Glacier exhibiting the greatest variation.

The newly determined sequences from 11 individuals across Alaska confirm previously published (Ortí *et al.*, 1994; Cresko, 2000; Weigner, 2012) locations of

restriction enzyme recognition sites. We did not observe any unexpected variation. These sequences have been deposited in GenBank (accession numbers KM508783-KM508793).

Geographical differences

Results of all statistical analyses on raw and resampled data were not significantly different. We therefore present only the results from the raw data. Coastline distance from Japan was a significant predictor of TNP relative abundance across the North Pacific without taking ecotype into account ($z = -18.62$, $P < 0.001$; Fig. A.3). This relationship also held true when considering only freshwater ($z = -15.09$, $P < 0.001$) or oceanic ($z = -3.16$, $P = 0.002$) individuals from Alaska, British Columbia and Oregon. Diagnostic plots of residuals from the overall and freshwater analyses revealed two major outliers: Barabara in south-central Alaska and Tashalich River in the Bering Glacier region (Fig. A.4). Barabara was the easternmost population fixed for the TNP haplotype group. The Tashalich River population also had a higher proportion of TNP individuals (51%) than other nearby populations. When considering only oceanic individuals, Middleton Island emerged as an outlier, having a greater representation of the TNP (19%) than other south-central Alaskan sites. Because model fits were comparable among freshwater, oceanic and pooled datasets, we performed a Mantel test to confirm significant effects of isolation by distance on only the pooled dataset. Our results were significant (Spearman's $r = 0.353$, $P < 0.001$), supporting the results of our linear model.

The role of deep-water trenches as impediments to migration

The Bussol' Strait represents the breakpoint between the mtDNA clade admixture zone to the north and the zone of TNP lineage fixation to the south (Fig. A.1). Similarly, the deep-water Amchitka Pass in the Aleutian Islands also appears to obstruct the movement of stickleback, because islands on either side of the pass are fixed for different clades. Statistical testing does not, however, support significant partitioning of genetic variation. AMOVA revealed that a significant amount of the variation was partitioned among individuals within populations (47.1%) and among populations grouped by location (40.9%; $P = 0.001$), and only 12% of the variation was partitioned across sites separated by trenches ($P = 0.174$).

Discussion

The North Pacific represents an extended zone of admixture between two divergent stickleback mtDNA clades. Mixed populations occur along a transect extending approximately 6000 km from Simushir in the Kuril Islands across Alaska to British Columbia, suggesting widespread movement of oceanic stickleback following Pleistocene deglaciation. Coastline distance from Japan is a significant predictor of the proportion of individuals carrying mtDNA of the TNP haplotype. At the basin-wide scale, the presence of regions where only one mtDNA clade is present adjacent to a broad zone of admixture suggests that dispersal from these eastern and western 'source' populations drives the observed patterns of clade distribution.

The distribution of the two mitochondrial lineages follows an overall pattern of isolation by distance at the basin and regional scales, but at small spatial scales, there is a great deal of mtDNA clade composition heterogeneity among sampled locations. In the

west, deep-water trenches in the Kuril and Aleutian island chains may impede stickleback migration (Fig. A.1). The Bussol' Strait represents the break-point between the mtDNA clade admixture zone to the north and the zone of TNP lineage fixation to the south (Fig. A.1) and has previously been identified as a biogeographical barrier influencing both plant and animal distributions (Barkalov, 2002; Bogatov, 2002; Kostenko, 2002; Lelej *et al.*, 2002; Teslenko, 2002; Pietsch *et al.*, 2003). The apparent absence of individuals carrying mtDNA from the ENA haplotype group south of the Bussol' Strait suggests that north to south movement across the strait is highly restricted or impossible.

Similarly, the deep-water Amchitka Pass in the Aleutian Islands appears to obstruct the movement of stickleback, because islands on either side of the pass are fixed for different clades. This portion of the Aleutian chain has previously been identified as a region of species turnover in skates (*Bathyraja*; Spies *et al.*, 2011), small Pacific Ocean perch (*Sebastes alutus*), Atka mackerel (*Pleurogrammus monopterygius*; Logerwell *et al.*, 2005) and rougheye rockfish (*Sebastes aleutianus*; Gharrett *et al.*, 2005). Our AMOVA results demonstrate, however, that no significant amount of variation is partitioned among groups of populations separated by either of these deep-water trenches, suggesting that mechanisms other than geomorphological barriers may drive the distribution of clade frequencies that we observe across these trenches.

Another potential explanation for the distribution of haplotypes along the Kuril archipelago may be ecological in nature. A species pair of three-spined stickleback is present in the Japanese archipelago and the Sea of Okhotsk – the Japan Sea stickleback and the Pacific Ocean stickleback (Kitano *et al.*, 2007, 2009). Despite being reproductively isolated due to genomic incompatibilities and divergent sex chromosomes, these two

species share mitochondrial haplotypes (Yamada *et al.*, 2001). Because these two species differ in ecology, behaviour and genetic architecture (Kitano *et al.*, 2009; Kume *et al.*, 2010), these factors could also explain the distribution of haplotypes in this region. Considering that Japan Sea stickleback are found on Hokkaido and extend into the Kuril Islands, it is possible that some of the discordance in haplotype distributions that we observe is due to ecological differences among habitats rather than, or in addition to, migration barriers caused by deep-water trenches.

In addition, Markov chain Monte Carlo approaches to modelling migration rates using sequences from Ortí *et al.* (1994) are consistent with asymmetrical rates of gene flow across the Pacific Basin. Integrated likelihood surfaces are consistent with a model suggesting ongoing migration from western to eastern populations, but little gene flow in the other direction (Nielsen & Wakeley, 2001). The distribution that we observe may also be due in part to differences in the timing of population expansions. For example, the TNP lineage may have spread more rapidly from Japan, which was not extensively glaciated, than the ENA clade from North America, which was largely glaciated until about 13,000 years ago.

The pattern of alternative fixation of the two mtDNA types suggests a dynamic demographic regime where founder effects, unstable population structure (e.g. stochastic history of colonization and extinction) and colonization barriers (Johnson & Taylor, 2004) may be the primary drivers of local population genetic diversity. Although previous research has not provided evidence of consistent links between mtDNA clade type and morphology (Deagle *et al.*, 1996; Johnson & Taylor, 2004) or nuclear genotype (Cresko, 2000), we cannot rule out a possible role for deterministic evolutionary processes (e.g.

selection or genomic co-adaptation) in shaping the observed patterns of mtDNA clade distribution.

This study presents the first densely sampled survey of three-spined stickleback mtDNA clade diversity across the North Pacific. We identified a region of clade admixture spanning approximately 6000 km of coastline from Simushir in the Kuril Islands to British Columbia, with a great deal of spatial heterogeneity in clade proportions among populations. We are the first to report the presence of the ENA haplotype west of south-central Alaska and our findings support previous work (Ortí *et al.*, 1994; Deagle *et al.*, 1996; Johnson & Taylor, 2004) that British Columbia represents the south-eastern extent of the TNP haplotype; the apparent absence of the TNP lineage south of Vancouver Island hints at a unidirectional barrier that impedes southward migrations.

The two extant mtDNA lineages are highly divergent and easily identified because they are separated by 18 nucleotide substitutions that have accumulated over an estimated 1 Myr of divergence (Ortí *et al.*, 1994). Although we were limited to testing for barriers to migration using only one marker, we would predict similar patterns in neutral nuclear markers (e.g. Haglund *et al.*, 1992; Higuchi & Goto, 1996; Peichel *et al.*, 2004; Colosimo *et al.*, 2005; Jones *et al.*, 2012). Loci that experience barriers to gene flow may, however, help us to identify locally adapted gene complexes. A valuable direction for future research would be to examine sequence variation in nuclear loci in populations spanning the North Pacific, to test for significant partitioning of genetic variation by haplotype. This hypothesis has previously been tested using a set of microsatellite markers (Cresko, 2000), with results supporting regular gene flow between individuals belonging to the two lineages. However, even if individuals containing mtDNA from the two different clades freely

interbreed, there may be residual influence of this deep divergence on contemporary patterns of nuclear genetic variation at some loci, leading to patterns of cytonuclear disequilibrium, which can be detected using modern genotyping techniques.

Acknowledgements

We thank F. Mueter, R. Lucas, M. Currey and S. Vanderzwan for assistance in the development of this study. Logistical support was provided by the Alaska Maritime National Wildlife Refuge of the US Fish and Wildlife Service, especially from J. Williams, H. Renner and L. Spitler. Funding was provided by the University of Alaska Anchorage Graduate Student Association Scholarships, an NSF EPSCoR Landscape Genetics Grant, a University of Alaska Center for Global Change and Arctic Systems Research grant and funds from University of Alaska and LGL Limited to E.A.L. Funding was also provided by NSF grants DEB 0919234 to F.A.v.H. and DEB 0963767 to J.A.L. All procedures were approved by the University of Alaska Anchorage IACUC (protocols 2001vonHi1, 2004vonHi1, 2007vonHi1, 159870). Stickleback were collected in accordance with Alaska Department of Fish and Game permit numbers SF2002-002, SF2006-017, SF2008-059, SF2009-016, SF2009-038, SF2009-065, SF2010-028, SF2010-029, SF2010-030, SF2010-111, SF2011-167 and SF2011-067 and Oregon Department of Fish and Game permit number 16933.

References

Atcheson, M.E., Myers, K.W., Beauchamp, D.A. & Mantua, N.J. (2012a) Bioenergetic response by steelhead to variation in diet, thermal habitat, and climate in the North Pacific Ocean. *Transactions of the American Fisheries Society*, **141**, 1081–1096.

Atcheson, M.E., Myers, K.W., Davis, N.J. & Mantua, N.J. (2012b) Potential trophodynamic and environmental drivers of steelhead (*Oncorhynchus mykiss*) productivity in the North Pacific Ocean. *Fisheries Oceanography*, **21**, 321–335.

Barkalov, V.Y. (2002) Outline of vegetation. *Flora and fauna of the Kuril Islands: materials of the International Kuril Island Project* (ed. by S.Y. Storozhenko, V.V. Bogatov and A.S. Lelej), pp. 35–66. Dalnauka, Vladivostok.

Bell, M.A. & Foster, S.A. (1994) *The evolutionary biology of the threespine stickleback*. Oxford University Press, Oxford.

Bogatov, V.V. (2002) Biogeographical problems of the Kuril Archipelago. *Flora and fauna of the Kuril Islands: materials of the International Kuril Island Project* (ed. by S.Y. Storozhenko, V.V. Bogatov and A.S. Lelej), pp. 150–160. Dalnauka, Vladivostok.

Colosimo, P.F., Peichel, C.L., Nereng, K., Blackman, B.K., Shapiro, M.D., Schluter, D. & Kingsley, D.M. (2004) The genetic architecture of parallel armor plate reduction in threespine sticklebacks. *PLoS Biology*, **2**, e109.

Colosimo, P.F., Hosemann, K.E., Balabhadra, S., Villarreal, G., Jr, Dickson, M., Grimwood, J., Schmutz, J., Myers, R.M., Schluter, D. & Kingsley, D.M. (2005) Widespread parallel evolution in sticklebacks by repeated fixation of ectodysplasin alleles. *Science*, **307**, 1928–1933.

Cresko W.A. (2000) *The ecology and geography of speciation: a case study using an adaptive radiation of threespine stickleback in Alaska*. PhD Thesis, Clark University, Worcester, MA.

Cresko, W.A., Amores, A., Wilson, C., Murphy, J., Currey, M., Phillips, P., Bell, M.A., Kimmel, C.B. & Postlethwait, J.H. (2004) Parallel genetic basis for repeated evolution of armor loss in Alaskan threespine stickleback populations. *Proceedings of the National Academy of Sciences USA*, **101**, 6050–6055.

Deagle B.E., Reimchen, T.E. & Levin, D.B. (1996) Origins of endemic stickleback from the Queen Charlotte Islands: mitochondrial and morphological evidence. *Canadian Journal of Zoology*, **74**, 1045–1056.

Dionne, M., Caron, F., Dodson, J.J. & Bernatchez, L. (2008) Landscape genetics and hierarchical genetic structure in Atlantic salmon: the interaction of gene flow and local adaptation. *Molecular Ecology*, **17**, 2382–2396.

Excoffier, L., Smouse, P.E. & Quattro, J.M. (1992) Analysis of molecular variance inferred from metric distances among DNA haplotypes: application to human mitochondrial DNA restriction data. *Genetics*, **131**, 479–491.

Gach, M.H. & Reimchen, T.E. (1989) Mitochondrial DNA patterns among endemic stickleback from the Queen Charlotte Islands: a preliminary survey. *Canadian Journal of Zoology*, **67**, 1324–1328.

Gharrett, A.J., Matala, A.P., Peterson, E.L., Gray, A.K. & Heifetz, J. (2005) Two genetically distinct forms of rougheye rockfish (*Sebastes aleutianus*) are different species. *Transactions of the American Fisheries Society*, **134**, 242–250.

Haglund, T.R., Buth, D.G. & Lawson R. (1992) Allozyme variation and phylogenetic relationships of Asian, North American, and European populations of the threespine stickleback, *Gasterosteus aculeatus*. *Copeia*, **1992**, 432–443.

Higuchi, M. & Goto, A. (1996) Genetic evidence supporting the existence of two distinct species in the genus *Gasterosteus* around Japan. *Environmental Biology of Fishes*, **47**, 1–16.

von Hippel, F.A. (ed.) (2010) *Tinbergen's legacy in behaviour: sixty years of landmark stickleback papers*. Brill, Leiden.

Johnson, L.S. & Taylor, E.B. (2004) The distribution of divergent mitochondrial DNA lineages of threespine stickleback (*Gasterosteus aculeatus*) in the northeastern Pacific Basin: post-glacial dispersal and lake accessibility. *Journal of Biogeography*, **31**, 1073–1083.

Jones, F.C., Grabherr, M.G., Chan, Y-G.F. *et al.* (2012) The genomic basis of adaptive evolution in threespine sticklebacks. *Nature*, **484**, 55–61.

Jørgensen, H.B.H., Hansen, M.M., Bekkevold, D., Ruzzante, D.E., Loeschcke, V. (2005) Marine landscapes and population genetic structure of herring (*Clupea harengus* L.) in the Baltic Sea. *Molecular Ecology*, **14**, 3219–3234.

Kitano, J., Mori, S. & Peichel, C.L. (2007) Sexual dimorphism in the external morphology of threespine stickleback (*Gasterosteus aculeatus*). *Copeia*, **2007**, 336–349.

Kitano, J., Ross, J.A., Mori, S., Kume, M., Jones, F.C., Chan, Y-G.F., Absher, D.M., Grimwood, J., Schmutz, J., Myers, R.M., Kingsley, D.M. & Peichel, C.L. (2009) A role for a neo-sex chromosome in stickleback speciation. *Nature*, **461**, 1079–1083.

Kostenko, V.A. (2002) The terrestrial mammals. *Flora and fauna of the Kuril Islands: materials of the International Kuril Island Project* (ed. by S.Y. Storozhenko, V.V. Bogatov and A.S. Lelej), pp. 135–143. Dalnauka, Vladivostok.

Kume, M., Kitano, J., Mori, S. & Shibuya, T. (2010) Ecological divergence and habitat isolation between two migratory forms of Japanese threespine stickleback (*Gasterosteus aculeatus*). *Journal of Evolutionary Biology*, **23**, 1436–1446.

Lelej, A.S., Storozhenko, S.Y. & Kholin, S.K. (2002) The insects (Insecta). *Flora and fauna of the Kuril Islands: materials of the International Kuril Island Project* (ed. by S.Y. Storozhenko, V.V. Bogatov and A.S. Lelej), pp. 96–108. Dalnauka, Vladivostok.

Logerwell, E.A., Aydin, K., Barbeaux, S., Brown, E., Conners, M.E., Lowe, S., Orr, J.W., Ortiz, I., Reuter, R. & Spencer, P. (2005) Geographic patterns in the demersal ichthyofauna of the Aleutian Islands. *Fisheries Oceanography*, **14**, 93–112.

Mäkinen, H.S., Cano, J.C. & Merilä, J. (2008) Identifying footprints of directional and balancing selection in marine and freshwater three-spined stickleback (*Gasterosteus aculeatus*) populations. *Molecular Ecology*, **17**, 3565–3582.

Meirmans, P.G. & van Tienderen, P.H. (2004) GENOTYPE and GENODIVE: two programs for the analysis of genetic diversity of asexual organisms. *Molecular Ecology Notes*, **4**, 792–794.

Morita, K., Morita, S.H. & Fukuwaka, M. (2009) *Offshore distributions of anadromous lamprey and threespine stickleback*. Document 1192. North Pacific Anadromous Fish Commission, Vancouver. Available at: http://www.npafc.org/publications/Documents/PDF%202009/1192(Japan).pdf.

Nielsen, R. & Wakeley, J. (2001) Distinguishing migration from isolation: a Markov chain Monte Carlo approach. *Genetics*, **158**, 885–896.

O'Reilly, P., Reimchen, T.E., Beech, R. & Strobeck, C. (1993) Mitochondrial DNA in *Gasterosteus* and Pleistocene glacial refugium on the Queen Charlotte Islands, British Columbia. *Evolution*, **47**, 678–684.

Ortí, G., Bell, M.A., Reimchen, T.E. & Meyer, A. (1994) Global survey of mitochondrial DNA sequences in the threespine stickleback: evidence for recent migrations. *Evolution*, **48**, 608–622.

Östlund-Nilsson, S., Mayer, S.I., & Huntingford, F.A. (2006) *Biology of the three-spined stickleback*. Taylor & Francis, Boca Raton, FL.

Peichel, C.L., Nereng, K.S., Ohgi, K.A., Cole, B.L., Colosimo, P.F., Buerkle, C.A., Schluter, D. & Kingsley, D.M. (2001) The genetic architecture of divergence between threespine stickleback species. *Nature*, **414**, 901–905.

Peichel, C.L., Ross, J.A., Matson, C.K., Dickson, M., Grimwood, J., Schmutz, J., Myers, R., Mori, S., Schluter, D. & Kingsley, D.M. (2004) The master sex-determination locus in threespine sticklebacks is on a nascent Y chromosome. *Current Biology*, **14**, 1416–1424.

Pietsch, T.W., Bogatov, V.V., Amaoka, K., Zhuravlev, Y.N., Barkalov, V.Y., Gage, S., Takahashi, H., Lelej, A.S., Storozhenko, S.Y., Minakawa, N., Bennett, D.J., Anderson, T.R., Ôhara, M., Prozorova, L.A., Kuwahara, Y., Kholin, S.K., Yabe, M., Stevenson, D.E. & MacDonald, E.L. (2003) Biodiversity and biogeography of the islands of the Kuril Archipelago. *Journal of Biogeography*, **30**, 1297–1310.

Quinn, T.P. & Light, J.T. (1989) Occurrence of threespine sticklebacks (*Gasterosteus aculeatus*) in the open North Pacific Ocean: migration or drift? *Canadian Journal of Zoology*, **67**, 2850–2852.

R Core Team (2014). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Spies, I.B., Stevenson, D.E., Orr, J.W. & Hoff, G.R. (2011) Molecular systematics of the skate subgenus *Arctoraja* (*Bathyraja*: Rajidae) and support for an undescribed species, the leopard skate, with comments on the phylogenetics of *Bathyraja*. *Ichthyology Research,* **58**, 77–83.

Storfer, A., Murphy, M.A., Evans, J.S., Goldberg, C.S., Robinson, S., Spear, S.F., Dezzani, R., Delmelle, E., Vierling, L. & Waits, L.P. (2006) Putting the 'landscape' in landscape genetics. *Heredity*, **98**, 128–142.

Teslenko, V.A. (2002) The aquatic insects. *Flora and fauna of the Kuril Islands: materials of the International Kuril Island Project* (ed. by S.Y. Storozhenko, V.V. Bogatov and A.S. Lelej), pp. 109–117. Dalnauka, Vladivostok.

Weigner, H.L. (2012) *Freshwater fish biogeography in the Bering Glacier region, Alaska*. PhD Thesis, University of Alaska Fairbanks, Fairbanks, AK.

Yamada, M., Higuchi, M. & Goto, A. (2001) Extensive introgression of mitochondrial DNA found between two genetically divergent forms of threespine stickleback, *Gasterosteus aculeatus*, around Japan. *Environmental Biology of Fishes*, **61**, 269–284.

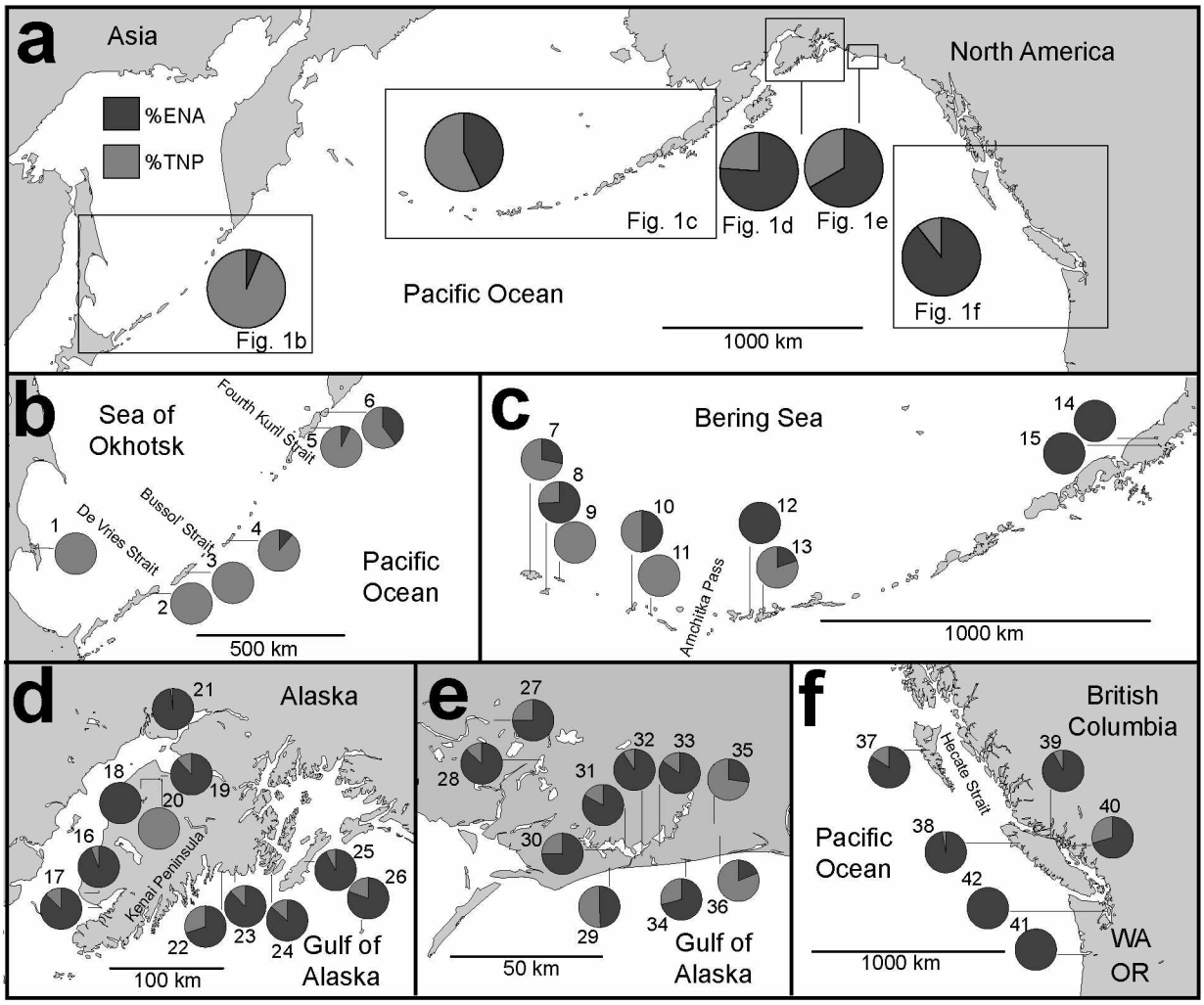

Figure A.1. Sampling locations of the three-spined stickleback (*Gasterosteus aculeatus*) with corresponding pie charts representing clade proportions. (a) North Pacific coast; (b) Kuril island chain; (c) Aleutian island chain and Alaska Peninsula; (d) south-central Alaska; (e) Bering Glacier forelands; and (f) west coast of North America. Numbers of locations correspond to those given in Table A.1.

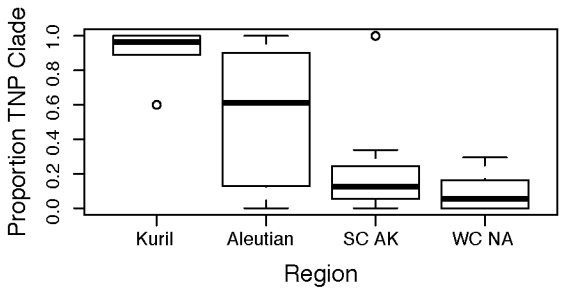

Figure A.2. Variation in trans-North-Pacific (TNP) clade proportions in the three-spined stickleback (*Gasterosteus aculeatus*) in the Kuril Islands (Kuril), Aleutian Islands (Aleutian), south-central Alaska (SC AK) and the west coast of North America (WC NA). Boxplots represent the median, quantiles, and outliers for each region.

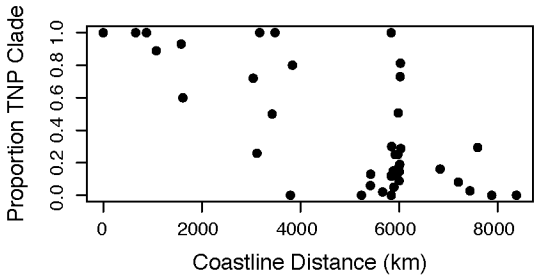

Figure A.3. Plot of the proportion of individuals of the three-spined stickleback (*Gasterosteus aculeatus*) belonging to the trans-North-Pacific (TNP) clade against coastline distance (km) from the westernmost study population (Sakhalin). Each point represents one sampling location in the North Pacific Ocean.

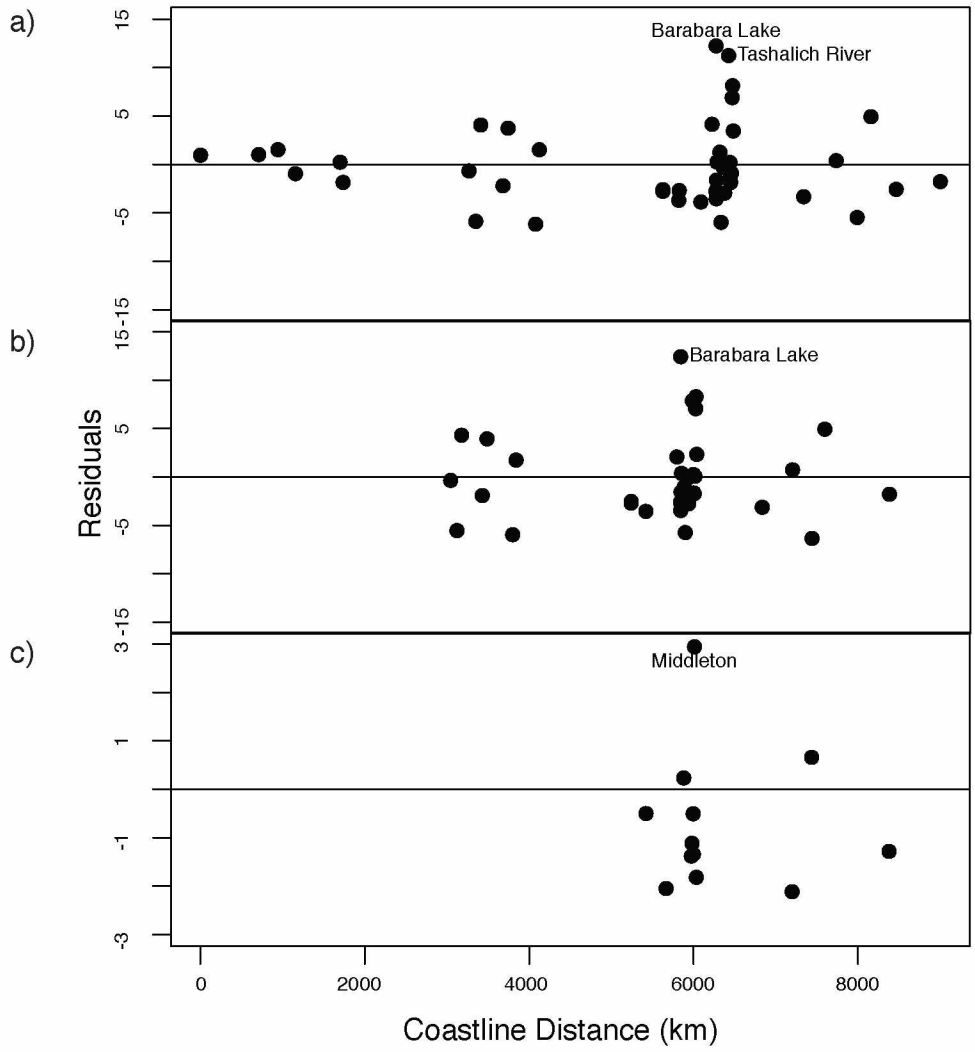

Figure A.4. Plots of residuals from the generalized linear model fit of coastline distance versus proportion of the trans-North-Pacific clade. Sites that are labeled were identified as outliers: (a) full data set; (b) freshwater individuals from Alaska, British Columbia and Oregon; and (c) oceanic individuals from Alaska, British Columbia and Oregon.

Table A.1. Sampling locations and their positions on Fig. A.1, number of sampling sites per location (n_S), number of individuals sampled per location (n_i) and percentages of European–North American clade (ENA) and trans-North-Pacific clade (TNP) individuals. Map position refers to the numbered locations in Fig.1. Sampling locations are listed from west to east. Data from Cook Inlet (Cresko, 2000; E.A.L., unpublished data), Bering Glacier (Weigner, 2012) and British Columbia (Deagle *et al.*, 1996; Johnson & Taylor, 2004) are from other studies.

Location	Map position	Latitude	Longitude	n_S	n_i	%ENA	%TNP
Kuril Islands	b				94	6	94
Sakhalin Island	1	46.7558	143.1286	3	16	0	100
Iturup Island	2	44.9878	147.7372	2	11	0	100
Urup Island	3	45.9542	150.1800	1	21	0	100
Simushir Island	4	47.0958	152.1375	1	27	11	89
Paramushir Island	5	50.7442	156.1550	1	14	7	93
Shumshu Island	6	50.7744	156.2594	2	5	40	60
Aleutian Islands + Alaska Peninsula	c				189	43	57
Attu Island	7	52.8908	172.9389	9	32	28	72
Agattu Island	8	52.4171	173.5943	9	31	74	26
Nizki Island	9	52.7363	173.9723	1	29	0	100
Kiska Island	10	51.9416	177.4988	4	24	50	50
Hawadax Island	11	51.8097	178.2297	1	20	0	100
Kanaga Island	12	51.7592	−177.2505	7	18	100	0
Adak Island	13	51.7669	−176.5902	1	20	20	80
Black Lake	14	56.4489	−158.9894	1	7	100	0
Chignik Island	15	56.2950	−158.4021	1	8	100	0
South-central Alaska	d				1238	78	22
Anchor River	16	59.7388	−151.7435	1	30	94	6
Homer Spit	17	59.6007	−151.4160	1	30	87	13
Sportfish Lake	18	60.8612	−150.2861	1	19	100	0
Two Island Lake	19	60.8827	−150.2739	1	17	88	12
Barabara Lake	20	60.8426	−150.2332	1	59	0	100
Rabbit Slough	21	61.5365	−149.2207	1	30	98	2
Horsehead Lake	22	59.9727	−149.0314	1	30	70	30
Salmonberry Pond	23	59.9681	−148.7991	1	50	88	12
Danger Island	24	59.9257	−148.0817	2	96	86	14
Montague Island	25	59.7848	−147.8686	5	157	92	8
Middleton Island	26	59.4390	−146.3309	14	720	81	19
Bering Glacier forelands	e				688	66	34
Pond near Bering River	27	60.4550	−144.2080	1	48	75	25
Bering River Lake	28	60.3600	−143.9730	1	83	87	13
Tashalich River	29	60.0390	−143.6300	1	79	49	51
Unnamed pond	30	60.1020	−143.5500	1	96	75	25
Vitus Lake	31	60.1254	−143.4171	1	86	83	17
Creek 1	32	60.1030	−143.3600	1	79	91	9
Pond near Vitus Lake	33	60.1159	−143.2926	1	48	85	15

Table A.2 continued.

Tsiu River	34	60.0800	−143.0500	1	73	71	29
Quonset Hut Lake	35	60.1679	−142.9870	1	48	27	73
East Tsivat Lake	36	60.1351	−142.9760	1	48	19	81
West coast of North America	f				1473	92	8
Queen Charlotte Islands	37	52.6438	−131.8079	46	688	89	11
Vancouver Island	38	49.6995	−125.6799	25	550	98	2
Central Coast	39	50.3963	−125.6369	7	89	99	1
Quadra Island	40	50.1903	−125.2652	4	75	71	29
Oregon	41	43.9890	−124.1250	3	29	100	0
Washington	42	47.6417	−122.8271	1	42	100	0
Total					3682	78	22